

Rec'd PCT/PTO 24 FEB 2005

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



525674

(43) Internationales Veröffentlichungsdatum
25. März 2004 (25.03.2004)

PCT

(10) Internationale Veröffentlichungsnummer
WO 2004/024932 A2

(51) Internationale Patentklassifikation⁷: C12P 13/04,
13/12

(21) Internationales Aktenzeichen: PCT/EP2003/009452

(22) Internationales Anmeldedatum:
26. August 2003 (26.08.2003)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:
102 39 073.8 26. August 2002 (26.08.2002) DE

(71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von US): BASF AKTIENGESELLSCHAFT [DE/DE];
67056 Ludwigshafen (DE).

(72) Erfinder; und
(75) Erfinder/Anmelder (nur für US): KRÖGER, Burkhard [DE/DE]; Im Waldhof 1, 67117 Limburgerhof (DE).
ZELDER, Oskar [DE/DE]; Franz-Stützel-Str. 8, 67346 Speyer (DE). KLOPPROGGE, Corinna [DE/DE]; Rastatter Str. 10, 68239 Mannheim (DE). SCHRÖDER, Hartwig [DE/DE]; Benzstr. 4, 69226 Nussloch (DE). HÄFNER, Stefan [DE/DE]; Luitpoldstr. 11, 67063 Ludwigshafen (DE).

(74) Anwalt: KINZEBACH, Werner; Reitstötter, Kinzebach & Partner (GbR), Sternwartstr. 4, 81679 München (DE).

(81) Bestimmungsstaaten (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR,

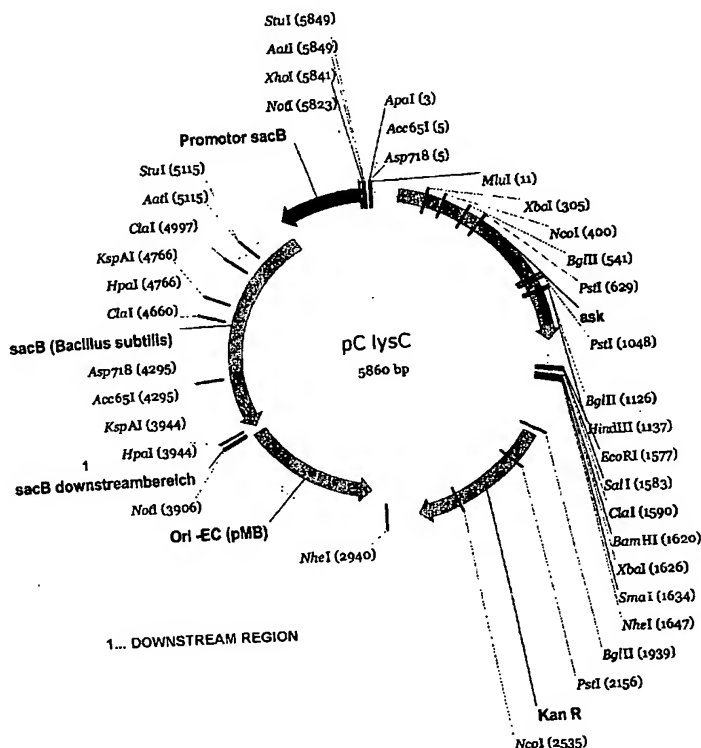
[Fortsetzung auf der nächsten Seite]

(54) Title: METHOD FOR ZYMOTIC PRODUCTION OF FINE CHEMICALS (META) CONTAINING SULPHUR

(54) Bezeichnung: VERFAHREN ZUR FERMENTATIVEN HERSTELLUNG SCHWEFELHALTIGER FEINCHEMIKALIEN (META)

(57) Abstract: The invention relates to methods for the zymotic production of fine chemicals, especially L-methionine, containing sulphur using bacteria, wherein a nucleotide sequence coding for a methionine-synthase (methA)-gene is expressed.

(57) Zusammenfassung: Die Erfindung betrifft Verfahren zur fermentativen Herstellung von schwefelhaltigen Feinchemikalien, insbesondere L-Methionin, unter Verwendung von Bakterien, in denen eine für ein Methionin-Synthase (methA)-Gen kodierende Nukleotidsequenz exprimiert wird.



Best Available Copy

WO 2004/024932 A2



KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

PT, RO, SE, SI, SK, TR), OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht:

— ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts

(84) **Bestimmungsstaaten (regional):** ARIPO-Patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL,

Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

VERFAHREN ZUR FERMENTATIVEN HERSTELLUNG SCHWEFELHALTIGER FEINCHEMIKALIEN (META)

Beschreibung

Gegenstand der Erfindung ist ein Verfahren zur fermentativen Herstellung von schwefelhaltigen Feinchemikalien, insbesondere L-Methionin, unter Verwendung von Bakterien, in denen eine für
5 ein Homoserin-O-Acetyl-Transferase (metA)-Gen kodierende Nukleotidsequenz exprimiert wird.

Stand der Technik

10 Schwefelhaltige Feinchemikalien, wie zum Beispiel Methionin, Homocystein, S-Adenosyl-Methionin, Glutathion, Cystein, Biotin, Thiamin, Liponsäure werden über natürliche Stoffwechselprozesse in Zellen hergestellt und werden in vielen Industriezweigen verwendet, einschließlich der Nahrungsmittel-, Futtermittel-, Kosmetik- und pharmazeutischen Industrie. Diese Substanzen, die zusammen als "schwefelhaltige Feinchemikalien" bezeichnet werden, umfassen organische Säuren, sowohl proteinogene als auch nicht-proteinogene Aminosäuren, Vitamine
15 und Cofaktoren. Ihre Produktion erfolgt am zweckmäßigsten im Großmaßstab mittels Anzucht von Bakterien, die entwickelt wurden, um große Mengen der jeweils gewünschten Substanz zu produzieren und sezernieren. Für diesen Zweck besonders geeignete Organismen sind coryneforme Bakterien, gram-positive nicht-pathogene Bakterien.

20 Es ist bekannt, dass Aminosäuren durch Fermentation von Stämmen coryneformer Bakterien, insbesondere Corynebacterium glutamicum, hergestellt werden. Wegen der großen Bedeutung wird ständig an der Verbesserung der Herstellverfahren gearbeitet. Verfahrensverbesserungen können fermentationstechnische Maßnahmen, wie zum Beispiel Rührung und Versorgung mit Sauerstoff, oder die Zusammensetzung der Nährmedien, wie zum Beispiel die Zuckerkonzentration während der Fermentation, oder die Aufarbeitung zum Produkt, beispielsweise durch Ionenaustauschchromatographie, oder die intrinsischen Leistungseigenschaften des Mikroorganismus selbst betreffen.
25

Über Stammselektion sind eine Reihe von Mutantenstämmen entwickelt worden, die ein Sortiment wünschenswerter Verbindungen aus der Reihe der schwefelhaltigen Feinchemikalien produzieren. Zur Verbesserung der Leistungseigenschaften dieser Mikroorganismen hinsichtlich der Produktion eines bestimmten Moleküls werden Methoden der Mutagenese, Selektion und Mutantenauswahl angewendet. Dies ist jedoch ein zeitaufwendiges und schwieriges Verfahren. Auf diese Weise erhält man z.B. Stämme, die resistent gegen Antimetabolite, wie z. B. die Methionin-Analoga α -Methyl-Methionin, Ethionin, Norleucin, N-Acetylnorleucin, S-Trifluoromethylhomocystein, 2-Amino-5-heptenoisäure, Seleno-Methionin, Methioninsulfoximin,
30
35

Methoxin, 1-Aminocyclopentan-Carboxylsäure oder auxotroph für regulatorisch bedeutsame Metabolite sind und schwefelhaltige Feinchemikalien, wie z. B. L-Methionin, produzieren.

Seit einigen Jahren werden ebenfalls Methoden der rekombinanten DNA-Technik zur Stammverbesserung von L-Aminosäure produzierender Stämme von *Corynebacterium* eingesetzt, indem man einzelne Aminosäure-Biosynthesegene amplifiziert und die Auswirkung auf die Aminosäure-Produktion untersucht.

Kurze Beschreibung der Erfindung

Der Erfindung lag die Aufgabe zugrunde, ein neues Verfahren zur verbesserten fermentativen Herstellung von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, bereitzustellen.

Gelöst wird obige Aufgabe durch Bereitstellung eines Verfahrens zur fermentativen Herstellung einer schwefelhaltigen Feinchemikalie, umfassend die Expression einer heterologen Nukleotidsequenz, welche für ein Protein mit metA-Aktivität kodiert, in einem coryneformen Bakterium.

Ein erster Gegenstand der Erfindung ist Verfahren zur fermentativen Herstellung wenigstens einer schwefelhaltigen Feinchemikalie, welches folgende Schritte umfasst:

- a) Fermentation einer die gewünschte schwefelhaltige Feinchemikalie produzierenden coryneformen Bakterienkultur, wobei in den coryneformen Bakterien zumindest eine heterologe Nukleotidsequenz exprimiert wird, welche für ein Protein mit Homoserin-O-Acetyl-Transferase (metA) –Aktivität kodiert;
- b) Anreicherung der schwefelhaltigen Feinchemikalie im Medium oder in den Zellen der Bakterien, und
- c) Isolieren der schwefelhaltigen Feinchemikalie, welche vorzugsweise L-Methionin umfasst.

Vorzugsweise besitzt obige heterologe metA-kodierende Nukleotidsequenz zur metA-kodierenden Sequenz aus *Corynebacterium glutamicum* ATCC 13032 eine Sequenzhomologie von weniger als 100% und vorzugsweise von mehr als 70% aufweist. Die metA-kodierende Sequenz ist vorzugsweise aus einem der folgenden Organismen von Liste I abgeleitet:

Liste I

<i>Corynebacterium diptheriae</i>	ATCC 14779
<i>Mycobacterium leprae</i>	ATCC 43910
<i>Mycobacterium tuberculosis</i> CDC1551	ATCC 25584
<i>Chlorobium tepidum</i>	ATCC 49652
<i>Pseudomonas aeruginosa</i>	ATCC 17933
<i>Caulobacter crescentus</i>	ATCC 19089
<i>Neisseria gonorrhoeae</i>	ATCC 53420
<i>Neisseria meningitidis</i>	ATCC 53414
<i>Pseudomonas fluorescens</i>	ATCC 13525
<i>Burkholderia cepacia</i>	ATCC 25416
<i>Nitrosomonas europaea</i>	ATCC 19718
<i>Haemophilus influenzae</i>	ATCC 51907
<i>Halobacterium</i> sp NRC1	ATCC 33170
<i>Thermus thermophilus</i>	ATCC 27634
<i>Deinococcus radiodurans</i>	ATCC 13939
<i>Saccharomyces cerevisiae</i>	ATCC 10751
<i>Schizosaccharomyces pombe</i>	ATCC 24969
<i>Xylella fastidiosa</i>	ATCC 35881
<i>Emericella nidulans</i>	ATCC 36104
<i>Mesorhizobium loti</i>	ATCC 35173
<i>Acremonium crysogenum</i>	ATCC 11550
<i>Pseudomonas putida</i>	ATCC 47054
<i>Staphylococcus aureus</i>	ATCC 35556

ATCC: American Type Culture Collection, Rockville, MD, USA

- 5 Die erfindungsgemäß eingesetzte metA-kodierende Sequenz umfasst vorzugsweise eine kodierende Sequenz gemäß SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43 und 45 oder eine dazu homologe Nukleotidsequenz, welche für ein Protein mit metA-Aktivität kodiert.
- 10 Die erfindungsgemäß eingesetzte metA-kodierende Sequenz kodiert außerdem vorzugsweise für ein Protein mit metA-Aktivität, wobei das Protein eine Aminosäuresequenz gemäß SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44 und 46 oder eine dazu homologe Aminosäuresequenz, welche für ein Protein mit metA-Aktivität steht, umfasst.
- 15 Die kodierende metA-Sequenz ist vorzugsweise eine in coryneformen Bakterien replizierbare oder eine stabil in das Chromosom integrierte DNA oder eine RNA.
- Gemäß einer bevorzugten Ausführungsform wird das erfindungsgemäße Verfahren durchgeführt, indem man
- 20

- a) einen mit einem Plasmidvektor transformierten Bakterienstamm einsetzt der wenigstens eine Kopie der kodierenden metA-Sequenz unter der Kontrolle regulativer Sequenzen trägt, oder
- b) einen Stamm einsetzt, in dem die kodierende metA-Sequenz in das Chromosom des Bakteriums integriert wurde

5

Es ist weiterhin bevorzugt, die kodierende metA-Sequenz für die Fermentation zu überexprimieren.

10

Außerdem kann es wünschenswert sein, Bakterien zu fermentieren, in denen zusätzlich wenigstens ein weiteres Gen des Biosyntheseweges der gewünschten schwefelhaltigen Feinchemikalie verstärkt ist; und / oder

in denen wenigstens ein Stoffwechselweg zumindest teilweise ausgeschaltet sind, der die Bildung der gewünschten schwefelhaltigen Feinchemikalie verringert.

15

Außerdem kann es wünschenswert sein, Bakterien zu fermentieren, in denen zusätzlich wenigstens ein weiteres Gen des Biosyntheseweges der gewünschten schwefelhaltigen Feinchemikalie durch Stoffwechselmetabolite in seiner Aktivität nicht in unerwünschter Weise beeinflusst wird.

20

Gemäß einer weiteren Ausführungsform des erfindungsgemäßen Verfahrens werden deshalb coryneforme Bakterien fermentiert, in denen gleichzeitig wenigstens eines der Gene, ausgewählt unter

25

- a) dem für eine Aspartatkinase kodierenden Gen lysC,
- b) dem für eine Aspartat-Semialdehyd-Dehydrogenase kodierenden Gen asd
- c) dem für die Glycerinaldehyd-3-Phosphat Dehydrogenase kodierenden Gen gap,
- d) dem für die 3-Phosphoglycerat Kinase kodierenden Gen pgk,
- e) dem für die Pyruvat Carboxylase kodierenden Gen pyc,
- f) dem für die Triosephosphat Isomerase kodierenden Gen tpi,
- g) dem für die Methionin Synthase kodierenden Gen methH,
- h) dem für die Cystathionin-gamma-Synthase kodierenden Gen metB,
- i) dem für die Cystathionin-gamma-Lyase kodierenden Gen metC,
- j) dem für die Serin-Hydroxymethyltransferase kodierenden Gen glyA,
- k) dem für die O-Acetylhomoserin-Sulphydrylase kodierenden Gen metY,
- l) dem für die Methylen-Tetrahydrofolat-Reduktase kodierenden Gen, metF
- m) dem für die Phosphoserin-Aminotransferase kodierenden Gen serC
- n) dem für die Phosphoserin-Phosphatase kodierenden Gen serB,

35

- o) dem für die Serine Acetyl-Transferase kodierenden Gen cysE,
 - p) dem für die Homoserin-Dehydrogenase kodierenden Gen hom,
- überexprimiert ist.

5 Gemäß einer anderen Ausführungsform des erfindungsgemäßen Verfahrens werden corynefor-
me Bakterien fermentiert, in denen gleichzeitig wenigstens eines der Gene ausgewählt unter
Genen der oben genannten Gruppe a) bis p) mutiert ist, so dass die korrespondierenden Protei-
ne, verglichen mit nicht mutierten Proteinen, in geringerem Maße oder nicht durch Stoffwech-
selmetabolite in ihrer Aktivität beeinflusst werden und dass insbesondere die erfindungsgemäße
10 Produktion der Feinchemikalie nicht beeinträchtigt wird.

Gemäß einer anderen Ausführungsform des erfindungsgemäßen Verfahrens werden corynefor-
me Bakterien fermentiert, in denen gleichzeitig wenigstens eines der Gene, ausgewählt unter

- q) dem für die Homoserine-Kinase kodierenden Gen thrB,
 - 15 r) dem für die Threonin Dehydratase kodierenden Gen ilvA,
 - s) dem für die Threonin Synthase kodierenden Gen thrC
 - t) dem für die Meso-Diaminopimelat D-Dehydrogenase kodierenden Gen ddh
 - u) dem für die Phosphoenolpyruvat-Carboxykinase kodierenden Gen pck,
 - v) dem für die Glucose-6-Phosphat-6-Isomerase kodierenden Gen pgj,
 - 20 w) dem für die Pyruvat-Oxidase kodierenden Gen poxB,
 - x) dem für die Dihydrodipicolinat Synthase kodierenden Gen dapA,
 - y) dem für die Dihydrodipicolinat Reduktase kodierenden Gen dapB; oder
 - z) dem für die Diaminopicolinat Decarboxylase kodierenden Gen lysA
- abschwächt ist, insbesondere durch Verringerung der Expressionsrate des korrespondierenden
25 Gens.

Gemäß einer anderen Ausführungsform des erfindungsgemäßen Verfahrens werden corynefor-
me Bakterien fermentiert, in denen gleichzeitig wenigstens eines der Gene der obigen Gruppen
q) bis z) mutiert ist, so dass die enzymatische Aktivität des korrespondierenden Proteins teilwei-
30 se oder vollständig verringert wird.

Vorzugsweise werden in dem erfindungsgemäßen Verfahren Mikroorganismen der Art Coryne-
bacterium glutamicum eingesetzt.

35 Ein weiterer Gegenstand der Erfindung betrifft ein Verfahren zur Herstellung eines L-Methionin-
haltigen Tierfuttermittel-Additivs aus Fermentationsbrühen, welches folgende Schritte umfasst

- a) Kultivierung und Fermentation eines L-Methionin produzierenden Mikroorganismus in einem Fermentationsmedium;
- b) Entfernung von Wasser aus der L-Methionin haltigen Fermentationsbrühe;
- c) Entfernung der während der Fermentation gebildeten Biomasse in einer Menge von 0 bis 100 Gew.-%; und
- d) Trocknung der gemäß b) und/oder c) erhaltenen Fermentationsbrühe, um das Tierfuttermittel-Additiv in der gewünschten Pulver- oder Granulatform zu erhalten.

10 Gegenstand der Erfindung sind ebenfalls die erstmalig aus obigen Mikroorganismen isolierten kodierenden metA-Sequenzen, die davon kodierten Homoserin-O-Acetyl-Transferase sowie die funktionalen Homologen dieser Polynukleotide bzw. Proteine.

Detaillierte Beschreibung der Erfindung

15 a) Allgemeine Begriffe

Als Proteine mit der Aktivität der Homoserin-O-Acetyl-Transferase auch metA (EC 2.3.1.31) genannt, werden solche Proteine beschrieben, die in der Lage sind Homoserin und Acetyl-Co-EnzymA zu O-Acetyl-Homoserin umzusetzen. Der Fachmann unterscheidet die Aktivität der Homoserin-O-Acetyl-Transferase von der Homoserin-O-Succinyl-Transferase, die in der Literatur aber auch metA genannt wird. In dem letztgenannten Enzym dient Succinyl-Coenzym A und nicht Acetyl-Coenzym A als Substrat der Reaktion. Der Fachmann kann die enzymatische Aktivität von Homoserin-O-Acetyl-Transferase durch Enzymtests nachweisen, Vorschriften dafür können sein: Park SD. Lee JY. Kim Y. Kim JH. Lee HS. Molecules & Cells. 8(3):286-94, 1998.

25 Im Rahmen der vorliegenden Erfindung umfasst der Begriff „schwefelhaltige Feinchemikalie“ jegliche chemische Verbindung, die wenigstens ein Schwefelatom kovalent gebunden enthält und durch ein erfindungsgemäßes Fermentationsverfahrens zugänglich ist. Nichtlimitierende Beispiele dafür sind Methionin, Homocystein, S-Adenosyl-Methionin, insbesondere Methio-
30 nin, und S-Adenosyl-Methionin.

Im Rahmen der vorliegenden Erfindung umfassen die Begriffe „L-Methionin“, „Methionin“, Homocystein und S-Adenosylmethionin auch die korrespondierenden Salze, wie z. B. Methionin-Hydrochlorid oder Methionin-Sulfat.

"Polynukleotide" bezeichnet im allgemeinen Polyribonukleotide (RNA) und Polydeoxyribonukleotide (DNA), wobei es sich um nicht modifizierte RNA oder DNA oder modifizierte RNA oder DNA handeln kann.

- 5 Unter "Polypeptiden" versteht man erfindungsgemäß Peptide oder Proteine, die zwei oder mehr über Peptidbindungen verbundene Aminosäuren enthalten.

Der Begriff „Stoffwechselmetabolit“ bezeichnet chemische Verbindungen, die im Stoffwechsel von Organismen als Zwischen- oder auch Endprodukte vorkommen und die neben ihrer Eigen-
10 schaft als chemische Bausteine auch modulierende Wirkung auf Enzyme und ihre katalytische Aktivität haben können. Dabei ist aus der Literatur bekannt, dass solche Stoffwechselmetabolite sowohl hemmend als auch stimulierend auf die Aktivität von Enzymen wirken können (Biochemistry, Stryer, Lubert, 1995 W. H. Freeman & Company, New York, New York.). In der Literatur ist auch beschrieben, dass es möglich ist durch Maßnahmen wie Mutation der genomischen
15 DNA durch UV-Strahlung, ionisierender Strahlung oder mutagene Substanzen und nachfolgender Selektion auf bestimmte Phänotypen in Organismen solche Enzyme zu produzieren, in denen die Beeinflussung durch Stoffwechselmetabolite verändert wurde (Sahm H. Eggeling L. de Graaf AA. Biological Chemistry 381(9-10):899-910, 2000; Eikmanns BJ. Eggeling L. Sahm H. Antonie van Leeuwenhoek. 64:145-63, 1993-94). Diese veränderten Eigenschaften können auch
20 durch gezielte Maßnahmen erreicht werden. Dabei ist dem Fachmann bekannt, dass in Genen für Enzyme bestimmte Nukleotide der für das Protein kodierenden DNA gezielt verändert werden können, so dass das aus der exprimierten DNA-Sequenz resultierende Protein bestimmte neue Eigenschaften aufweist, so zum Beispiel, dass die modulierende Wirkung von Stoffwechselmetaboliten gegenüber dem nicht veränderten Protein verändert ist

25 Die Begriffe "exprimieren" bzw. "Verstärkung" oder „Überexpression“ beschreiben im Kontext der Erfindung die Produktion bzw. Erhöhung der intrazellulären Aktivität eines oder mehrerer Enzyme in einem Mikroorganismus, die durch die entsprechende DNA kodiert werden. Dazu kann man beispielsweise ein Gen in einen Organismus einbringen, ein vorhandenes Gen durch ein
30 anderes Gen ersetzen, die Kopienzahl des Gens bzw. der Gene erhöhen, einen starken Promotor verwenden oder ein Gen verwenden, das für ein entsprechendes Enzym mit einer hohen Aktivität kodiert und man kann gegebenenfalls diese Maßnahmen kombinieren.

b) Erfindungsgemäße metA-Proteine

35 Erfindungsgemäß mit umfasst sind ebenfalls „funktionale Äquivalente“ der konkret offenbarten

metA-Enzyme aus Organismen obiger Liste I.

„Funktionale Äquivalente“ oder Analoga der konkret offenbarten Polypeptide sind im Rahmen der vorliegenden Erfindung davon verschiedene Polypeptide, welche weiterhin die gewünschte biologische Aktivität, wie z.B. Substratspezifität, besitzen.

5

Unter "funktionalen Äquivalenten" versteht man erfindungsgemäß insbesondere Mutanten, welche in wenigstens einer der oben genannten Sequenzpositionen eine andere als die konkret genannte Aminosäure aufweisen aber trotzdem eine der oben genannten biologische Aktivitäten besitzen. "Funktionale Äquivalente" umfassen somit die durch eine oder mehrere Aminosäure-Additionen, -Substitutionen, -Deletionen und/oder -Inversionen erhältlichen Mutanten, wobei die genannten Veränderungen in jeglicher Sequenzposition auftreten können, solange sie zu einer Mutante mit dem erfindungsgemäßen Eigenschaftsprofil führen. Funktionale Äquivalenz ist insbesondere auch dann gegeben, wenn die Reaktivitätsmuster zwischen Mutante und unverändertem Polypeptid qualitativ übereinstimmen, d.h. beispielsweise gleiche Substrate mit unterschiedlicher Geschwindigkeit umgesetzt werden.

10

15

"Funktionale Äquivalente" umfassen natürlich auch Polypeptide welche aus anderen Organismen zugänglich sind, sowie natürlich vorkommende Varianten. Beispielsweise lassen sich durch Sequenzvergleich Bereiche homologer Sequenzregionen festlegen und in Anlehnung an die konkreten Vorgaben der Erfindung äquivalente Enzyme ermitteln.

20

„Funktionale Äquivalente“ umfassen ebenfalls Fragmente, vorzugsweise einzelne Domänen oder der Sequenzmotive, der erfindungsgemäßen Polypeptide, welche z.B. die gewünschte biologische Funktion aufweisen.

25

„Funktionale Äquivalente“ sind außerdem Fusionsproteine, welche ein der oben genannten Polypeptidsequenzen oder davon abgeleitete funktionale Äquivalente und wenigstens eine weitere, davon funktionell verschiedene, heterologe Sequenz in funktioneller N- oder C-terminaler Verknüpfung (d.h. ohne gegenseitigen wesentliche funktionelle Beeinträchtigung der Fusionsproteinteile) aufweisen. Nichtlimitierende Beispiele für derartige heterologe Sequenzen sind z.B. Signalpeptide, Enzyme, Immunoglobuline, Oberflächenantigene, Rezeptoren oder Rezeptorliganden.

30

Erfindungsgemäß mit umfasste „funktionale Äquivalente“ sind Homologe zu den konkret offenbarten Proteinen. Diese besitzen wenigstens 20%, oder etwa 30%, 40%, 50 %, vorzugsweise

35

wenigstens etwa 60 %, 65%, 70%, oder 75% ins besondere wenigsten 85 %, wie z.B. 90%, 95% oder 99%, Homologie zu einer der konkret offenbarten Sequenzen, berechnet nach dem Algorithmus von Pearson und Lipman, Proc. Natl. Acad. Sci. (USA) 85(8), 1988, 2444-2448.

- 5 Homologe der erfindungsgemäßen Proteine oder Polypeptide können durch Mutagenese erzeugt werden, z.B. durch Punktmutation oder Verkürzung des Proteins. Der Begriff "Homolog", wie er hier verwendet wird, betrifft eine variante Form des Proteins, die als Agonist oder Antagonist der Protein-Aktivität wirkt.
- 10 Homologe des erfindungsgemäßen Proteine können durch Screening kombinatorischer Banken von Mutanten, wie z.B. Verkürzungsmutanten, identifiziert werden. Beispielsweise kann eine variierte Bank von Protein-Varianten durch kombinatorische Mutagenese auf Nukleinsäureebene erzeugt werden, wie z.B. durch enzymatisches Ligieren eines Gemisches synthetischer Oligonukleotide. Es gibt eine Vielzahl von Verfahren, die zur Herstellung von Banken potentieller
- 15 Homologer aus einer degenerierten Oligonukleotidsequenz verwendet werden können. Die chemische Synthese einer degenerierten Gensequenz kann in einem DNA-Syntheseautomaten durchgeführt werden, und das synthetische Gen kann dann in einen geeigneten Expressionsvektor ligiert werden. Die Verwendung eines degenerierten Gensatzes ermöglicht die Bereitstellung sämtlicher Sequenzen in einem Gemisch, die den gewünschten Satz an potentiellen Proteinsequenzen codieren. Verfahren zur Synthese degenerierter Oligonukleotide sind dem Fachmann bekannt (z.B. Narang, S.A. (1983) Tetrahedron 39:3; Itakura et al. (1984) Annu. Rev. Biochem. 53:323; Itakura et al., (1984) Science 198:1056; Ike et al. (1983) Nucleic Acids Res. 11:477).
- 25 Zusätzlich können Banken von Fragmenten des Protein-Codons verwendet werden, um eine variierte Population von Protein-Fragmenten zum Screening und zur anschließenden Selektion von Homologen eines erfindungsgemäßen Proteins zu erzeugen. Bei einer Ausführungsform kann eine Bank von kodierenden Sequenzfragmenten durch Behandeln eines doppelsträngigen PCR-Fragmentes einer kodierenden Sequenz mit einer Nuklease unter Bedingungen, unter denen ein Nicking nur etwa einmal pro Molekül erfolgt, Denaturieren der doppelsträngigen DNA,
- 30 Renaturieren der DNA unter Bildung doppelsträngiger DNA, die Sense-/Antisense-Paare von verschiedenen genickten Produkten umfassen kann, Entfernen einzelsträngiger Abschnitte aus neu gebildeten Duplices durch Behandlung mit S1-Nuclease und Ligieren der resultierenden Fragmentbank in einen Expressionsvektor erzeugt werden. Durch dieses Verfahren kann eine
- 35 Expressionsbank hergeleitet werden, die N-terminale, C-terminale und interne Fragmente mit verschiedenen Größen des erfindungsgemäßen Proteins kodiert.

Im Stand der Technik sind mehrere Techniken zum Screening von Genprodukten kombinatorischer Banken, die durch Punktmutationen oder Verkürzung hergestellt worden sind, und zum Screening von DNA-Banken auf Genprodukte mit einer ausgewählten Eigenschaft bekannt. Diese Techniken lassen sich an das schnelle Screening der Genbanken anpassen, die durch kombinatorische Mutagenese erfindungsgemäßer Homologer erzeugt worden sind. Die am häufigsten verwendeten Techniken zum Screening großer Genbanken, die einer Analyse mit hohem Durchsatz unterliegen, umfassen das Klonieren der Genbank in replizierbare Expressionsvektoren, Transformieren der geeigneten Zellen mit der resultierenden Vektorenbank und Expressieren der kombinatorischen Gene unter Bedingungen, unter denen der Nachweis der gewünschten Aktivität die Isolation des Vektors, der das Gen codiert, dessen Produkt nachgewiesen wurde, erleichtert. Recursive-Ensemble-Mutagenese (REM), eine Technik, die die Häufigkeit funktioneller Mutanten in den Banken vergrößert, kann in Kombination mit den Screeningtests verwendet werden, um Homologe zu identifizieren (Arkin und Yourvan (1992) PNAS 89:7811-7815; Delgrave et al. (1993) Protein Engineering 6(3):327-331

c) Erfindungsgemäße Polynukleotide

Gegenstand der Erfindung sind ebenso Nukleinsäuresequenzen (einzel- und doppelsträngige DNA- und RNA-Sequenzen, wie z.B. cDNA und mRNA), kodierend für eines der obigen metA-Enzyme und deren funktionalen Äquivalenten, welche z.B. auch unter Verwendung künstlicher Nukleotidanaloga zugänglich sind.

Die Erfindung betrifft sowohl isolierte Nukleinsäuremoleküle, welche für erfindungsgemäße Polypeptide bzw. Proteine oder biologisch aktive Abschnitte davon kodieren, sowie Nukleinsäurefragmente, die z.B. zur Verwendung als Hybridisierungssonden oder Primer zur Identifizierung oder Amplifizierung von erfindungsgemäßer kodierenden Nukleinsäuren verwendet werden können.

Die erfindungsgemäßen Nukleinsäuremoleküle können zudem untranslatierte Sequenzen vom 3'- und/oder 5'-Ende des kodierenden Genbereichs enthalten

Ein "isoliertes" Nukleinsäuremolekül wird von anderen Nukleinsäuremolekülen abgetrennt, die in der natürlichen Quelle der Nukleinsäure zugegen sind und kann überdies im wesentlichen frei von anderem zellulären Material oder Kulturmedium sein, wenn es durch rekombinante Techni-

ken hergestellt wird, oder frei von chemischen Vorstufen oder anderen Chemikalien sein, wenn es chemisch synthetisiert wird.

- 5 Die Erfindung umfasst weiterhin die zu den konkret beschriebenen Nukleotidsequenzen komplementären Nukleinsäuremoleküle oder einen Abschnitt davon.

- 10 Die erfindungsgemäß Nukleotidsequenzen ermöglichen die Erzeugung von Sonden und Primern, die zur Identifizierung und/oder Klonierung von homologer Sequenzen in anderen Zelltypen und Organismen verwendbar sind. Solche Sonden bzw. Primer umfassen gewöhnlich einen Nukleotidsequenzbereich, der unter stringenten Bedingungen an mindestens etwa 12, vorzugsweise mindestens etwa 25, wie z.B. etwa 40, 50 oder 75 aufeinanderfolgende Nukleotide eines Sense-Stranges einer erfindungsgemäßen Nukleinsäuresequenz oder eines entsprechenden Antisense-Stranges hybridisiert.
- 15 Weitere erfindungsgemäße Nukleinsäuresequenzen sind abgeleitet von SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43 oder 45 und unterscheiden sich davon durch Addition, Substitution, Insertion oder Deletion einzelner oder mehrerer Nukleotide, kodieren aber weiterhin für Polypeptide mit dem gewünschten Eigenschaftsprofil. Dies können Polynukleotide sein, die zu obigen Sequenzen in mindestens etwa 50%, 55%, 60%, 65%, 70%, 20 80% oder 90%, vorzugsweise in mindestens etwa 95%, 96%, 97%, 98% oder 99% der Sequenzpositionen identisch sind.

- 25 Erfindungsgemäß umfasst sind auch solche Nukleinsäuresequenzen, die sogenannte stumme Mutationen umfassen oder entsprechend der Codon-Nutzung eines speziellen Ursprungs- oder Wirtsorganismus, im Vergleich zu einer konkret genannten Sequenz verändert sind, ebenso wie natürlich vorkommende Varianten, wie z.B. Spleißvarianten oder Allelvarianten, davon. Gegenstand sind ebenso durch konservative Nukleotidsubstitutionen (d.h. die betreffende Aminosäure wird durch eine Aminosäure gleicher Ladung, Größe, Polarität und/oder Löslichkeit ersetzt) erhaltliche Sequenzen.

- 30 Gegenstand der Erfindung sind auch die durch Sequenzpolymorphismen von den konkret offenbarten Nukleinsäuren abgeleiteten Moleküle. Diese genetischen Polymorphismen können zwischen Individuen innerhalb einer Population aufgrund der natürlichen Variation existieren. Diese natürlichen Variationen bewirken üblicherweise eine Varianz von 1 bis 5 % in der Nukleotidsequenz eines Gens.
- 35

Weiterhin umfasst die Erfindung auch Nukleinsäuresequenzen, welchen mit oben genannten kodierenden Sequenzen hybridisieren oder dazu komplementär sind. Diese Polynukleotide lassen sich bei Durchmusterung von genomischen oder cDNA-Banken auffinden und gegebenenfalls daraus mit geeigneten Primern mittels PCR vermehren und anschließend beispielsweise mit geeigneten Sonden isolieren. Eine weitere Möglichkeit bietet die Transformation geeigneter Mikroorganismen mit erfindungsgemäßen Polynukleotiden oder Vektoren, die Vermehrung der Mikroorganismen und damit der Polynukleotide und deren anschließende Isolierung. Darüber hinaus können erfindungsgemäße Polynukleotide auch auf chemischem Wege synthetisiert werden.

Unter der Eigenschaft, an Polynukleotide „hybridisieren“ zu können, versteht man die Fähigkeit eines Poly- oder Oligonukleotids unter stringenten Bedingungen an eine nahezu komplementäre Sequenz zu binden, während unter diesen Bedingungen unspezifische Bindungen zwischen nicht-komplementären Partnern unterbleiben. Dazu sollten die Sequenzen zu 70-100%, vorzugsweise zu 90-100%, komplementär sein. Die Eigenschaft komplementärer Sequenzen, spezifisch aneinander binden zu können, macht man sich beispielsweise in der Northern- oder Southern-Blot-Technik oder bei der Primerbindung in PCR oder RT-PCR zunutze. Üblicherweise werden dazu Oligonukleotide ab einer Länge von 30 Basenpaaren eingesetzt. Unter stringenten Bedingungen versteht man beispielsweise in der Northern-Blot-Technik die Verwendung einer 50 – 70 °C, vorzugsweise 60 – 65 °C warmen Waschlösung, beispielsweise 0,1x SSC-Puffer mit 0,1% SDS (20x SSC: 3M NaCl, 0,3M Na-Citrat, pH 7,0) zur Elution unspezifisch hybridisierter cDNA-Sonden oder Oligonukleotide. Dabei bleiben, wie oben erwähnt, nur in hohem Maße komplementäre Nukleinsäuren aneinander gebunden. Die Einstellung stringenter Bedingungen ist dem Fachmann bekannt und ist z.B. in Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. beschrieben.

d) Isolierung der kodierenden metA-Gene

Die für das Enzym Homoserin-O-Acetyl-Transferase codierenden metA-Gene aus den Organismen obiger Liste I sind in an sich bekannter Weise isolierbar.

Zur Isolierung der metA-Gene oder auch anderer Gene der Organismen aus obiger Liste I wird zunächst eine Genbank dieses Organismus in Escherichia coli (E. coli) angelegt. Das Anlegen von Genbanken ist in allgemein bekannten Lehrbüchern und Handbüchern ausführlich beschrieben. Als Beispiel seien das Lehrbuch von Winnacker: Gene und Klone, Eine Einführung in die Gentechnologie (Verlag Chemie, Weinheim, Deutschland, 1990), oder das Handbuch von

Sambrook et al.: Molecular Cloning, A Laboratory Manual (Cold Spring Harbor Laboratory Press, 1989) genannt. Eine sehr bekannte Genbank ist die des E. coli K-12 Stammes W3110, die von Kohara et al. (Cell 50, 495-508 (198)) in λ -Vektoren angelegt wurde.

- 5 Zur Herstellung einer Genbank von Organismen der Liste I in E. coli können Cosmide, wie der Cosmidvektor SuperCos I (Wahl et al., 1987, Proceedings of the National Academy of Sciences USA, 84: 2160-2164), aber auch Plasmide, wie pBR322 (BoliVal; Life Sciences, 25, 807-818 (1979)) oder pUC9 (Vieira et al., 1982, Gene, 19: 259-268), verwendet werden. Als Wirte eignen sich besonders solche E. coli Stämme, die restriktions- und rekombinationsdefekt sind. Ein Bei-
- 10 spiel hierfür ist der Stamm DH5 α mcr, der von Grant et al. (Proceedings of the National Academy of Sciences USA, 87 (1990) 4645-4649) beschrieben wurde. Die mit Hilfe von Cosmiden klonierten langen DNA-Fragmente können anschließend wiederum in gängige, für die Sequenzierung geeignete Vektoren subkloniert und anschließend sequenziert werden, so wie es z. B. bei Sanger et al. (proceedings of the National Academy of Sciences of the United States of America, 74:
- 15 5463-5467, 1977) beschrieben ist.

Die erhaltenen DNA-Sequenzen können dann mit bekannten Algorithmen bzw. Sequenzanalyse-Programmen, wie z. B. dem von Staden (Nucleic Acids Research 14, 217-232 (1986)), dem von Marck (Nucleic Acids Research 16, 1829-1836 (1988)) oder dem GCG-Programm von Butler

20 (Methods of Biochemical Analysis 39, 74-97 (1998)), untersucht werden.

Die für die metA-Gene kodierenden DNA-Sequenzen von Organismen gemäß obiger Liste I wurde gefunden. Insbesondere wurden DNA-Sequenzen gemäß SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43 und 45 gefunden. Weiterhin wurden

25 aus diesen vorliegenden DNA-Sequenzen mit den oben beschriebenen Methoden die Aminosäuresequenzen der entsprechenden Proteine abgeleitet. Durch SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44 und 46 sind die sich ergebenden Aminosäuresequenzen der metA-Genprodukte dargestellt.

30 Kodierende DNA-Sequenzen, die sich aus den Sequenzen gemäß SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43 und 45 durch die Degeneration des genetischen Kodes ergeben, sind ebenfalls Gegenstand der Erfindung. In gleicher Weise sind DNA-Sequenzen, die mit diesen Sequenzen oder davon abgeleiteten Sequenzteilen hybridisieren, Gegenstand der Erfindung.

- Anleitungen zur Identifizierung von DNA-Sequenzen mittels Hybridisierung findet der Fachmann unter anderem im Handbuch "The DIG System Users Guide für Filter Hybridization" der Firma Boehringer Mannheim GmbH (Mannheim, Deutschland, 1993) und bei Liebl et al. (International Journal of Systematic Bacteriology (1991) 41: 255-260). Anleitungen zur Amplifikation von DNA-Sequenzen mit Hilfe der Polymerase-Kettenreaktion (PCR) findet der Fachmann unter anderem im Handbuch von Gait: Oligonukleotide synthesis: A Practical Approach (IRL Press, Oxford, UK, 1984) und bei Newton und Graham: PCR (Spektrum Akademischer Verlag, Heidelberg, Deutschland, 1994).
- 5
- 10 Weiterhin ist bekannt, dass Änderungen am N- und/oder C- Terminus eines Proteins dessen Funktion nicht wesentlich beeinträchtigen oder sogar stabilisieren können. Angaben hierzu findet der Fachmann unter anderem bei Ben-Bassat et al. (Journal of Bacteriology 169: 751-757 (1987)), bei O'Regan et al. (Gene 77: 237-251 (1989)), bei Sahin-Toth et al. (Protein Sciences 3: 240-247 (1994)), bei Hochuli et al. (Biotechnology 6: 1321-1325 (1988)) und in bekannten
- 15 Lehrbüchern der Genetik und Molekularbiologie.

Aminosäuresequenzen, die sich in entsprechender Weise aus den SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44 und 46 ergeben, sind ebenfalls Bestandteil der Erfindung.

20

e) Erfindungsgemäß verwendete Wirtszellen

- Weitere Gegenstände der Erfindung betreffen als Wirtszelle dienende Mikroorganismen, insbesondere coryneforme Bakterien, die einen Vektor, insbesondere Pendelvektor oder Plasmidvektor, der wenigstens ein metA-Gen erfindungsgemäßer Definition trägt, enthalten oder in denen ein erfindungsgemäßes metA-Gen exprimiert bzw. verstärkt ist.
- 25

- Diese Mikroorganismen können schwefelhaltige Feinchemikalien, insbesondere L-Methionin, aus Glucose, Saccharose, Lactose, Fructose, Maltose, Melasse, Stärke, Cellulose oder aus Glycerin und Ethanol herstellen. Vorzugsweise sind dies coryneforme Bakterien, insbesondere der Gattung Corynebacterium. Aus der Gattung Corynebacterium ist insbesondere die Art Corynebacterium glutamicum zu nennen, die in der Fachwelt für ihre Fähigkeit bekannt ist, L-Aminosäuren zu produzieren.
- 30

- Als Beispiele für geeignete Stämme coryneformer Bakterien sind solche der Gattung Corynebacterium, insbesondere der Art Corynebacterium glutamicum (C. glutamicum), wie
- 35

- Corynebacterium glutamicum ATCC 13032,
Corynebacterium acetoglutamicum ATCC 15806,
Corynebacterium acetoacidophilum ATCC 13870,
Corynebacterium thermoaminogenes FERM BP-1539,
5 Corynebacterium melassecola ATCC 17965

oder

der Gattung Brevibacterium, wie

Brevibacterium flavum ATCC 14067

- 10 Brevibacterium lactofermentum ATCC 13869 und
Brevibacterium divaricatum ATCC 14020 zu nennen;
oder davon abgeleitete Stämme, wie
Corynebacterium glutamicum KFCC10065
Corynebacterium glutamicum ATCC21608

15

welche ebenfalls die gewünschte Feinchemikalie oder deren Vorstufe(n) produzieren.

Mit der Abkürzung KFCC ist die Korean Federation of Culture Collection gemeint, mit der Abkürzung ATCC die American type strain culture collection, mit der Abkürzung FERM BP die Sammlung des National institute of Bioscience and Human-Technology, Agency of Industrial Science
20 and Technology, Japan bezeichnet.

f) Durchführung der erfindungsgemäßen Fermentation

- 25 Erfindungsgemäß wurde festgestellt, dass coryneforme Bakterien nach Überexpression eines metA-Gens aus Organismen der Liste I in vorteilhafter Weise schwefelhaltige Feinchemikalien, insbesondere L-Methionin, produzieren.

- 30 Zur Erzielung einer Überexpression kann der Fachmann unterschiedliche Maßnahmen einzeln oder in Kombination ergreifen. So kann die Kopienzahl der entsprechenden Gene erhöht werden, oder es kann die Promotor- und Regulationsregion oder die Ribosomenbindungsstelle, die sich stromaufwärts des Strukturgens befindet, mutiert werden. In gleicher Weise wirken Expressionskassetten, die stromaufwärts des Strukturgens eingebaut werden. Durch induzierbare Promotoren ist es zusätzlich möglich, die Expression im Verlaufe der fermentativen L-Methionin-Produktion zu steigern. Durch Maßnahmen zur Verlängerung der Lebensdauer der mRNA wird
35 ebenfalls die Expression verbessert. Weiterhin wird durch Verhinderung des Abbaus des Enzymproteins ebenfalls die Enzymaktivität verstärkt. Die Gene oder Genkonstrukte können ent-

weder in Plasmiden mit unterschiedlicher Kopienzahl vorliegen oder im Chromosom integriert und amplifiziert sein. Alternativ kann weiterhin eine Überexpression der betreffenden Gene durch Veränderung der Medienzusammensetzung und Kulturführung erreicht werden.

- 5 Anleitungen hierzu findet der Fachmann unter anderem bei Martin et al. (Bionotechnology 5, 137-146 (1987)), bei Guerrero et al. (Gene 138, 35-41 (1994)), Tsuchiya und Morinaga (Bio/Technology 6, 428-430 (1988)), bei Eikmanns et al. (Gene 102, 93-98 (1991)), in der Europäischen Patentschrift 0472869, im US Patent 4,601,893, bei Schwarzer und Pühler (Biotechnology 9, 84-87 (1991)), bei Remscheid et al. (Applied and Environmental Microbiology 60, 126-132 (1994)), bei LaBarre et al. (Journal of Bacteriology 175, 1001-1007 (1993)), in der Patentanmeldung WO 96/15246, bei Malumbres et al. (Gene 134, 15-24 (1993)), in der japanischen Offenlegungsschrift JP-A-10-229891, bei Jensen und Hammer (Biotechnology and Bioengineering 58, 191-195 (1998)), bei Makrides (Microbiological Reviews 60 : 512-538 (1996) und in bekannten Lehrbüchern der Genetik und Molekularbiologie.
- 15 Gegenstand der Erfindung sind deshalb auch Expressionskonstrukte, enthaltend unter der genetischen Kontrolle regulativer Nukleinsäuresequenzen eine für ein erfindungsgemäßes Polypeptid kodierende Nukleinsäuresequenz; sowie Vektoren, umfassend wenigstens eines dieser Expressionskonstrukte. Vorzugsweise umfassen solche erfindungsgemäßen Konstrukte 5'-stromaufwärts von der jeweiligen kodierenden Sequenz einen Promotor und 3'-stromabwärts eine Terminatorsequenz sowie gegebenenfalls weitere übliche regulative Elemente, und zwar jeweils operativ verknüpft mit der kodierenden Sequenz. Unter einer „operativen Verknüpfung“ versteht man die sequentielle Anordnung von Promotor, kodierender Sequenz, Terminator und gegebenenfalls weiterer regulativer Elemente derart, dass jedes der regulativen Elemente seine Funktion bei der Expression der kodierenden Sequenz bestimmungsgemäß erfüllen kann. Beispiele für operativ verknüpfbare Sequenzen sind Aktivierungssequenzen sowie Enhancer und dergleichen. Weitere regulative Elemente umfassen selektierbare Marker, Amplifikationssignale, Replikationsursprünge und dergleichen. Geeignete regulatorische Sequenzen sind z.B. beschrieben in Goeddel, Gene Expression Technology: Methods in Enzymology 185, Academic Press, San Diego, CA (1990).
- 20
- 25
- 30

Zusätzlich zu den artifiziellen Regulationssequenzen kann die natürliche Regulationssequenz vor dem eigentlichen Strukturgen noch vorhanden sein. Durch genetische Veränderung kann diese natürliche Regulation gegebenenfalls ausgeschaltet und die Expression der Gene erhöht oder erniedrigt werden. Das Genkonstrukt kann aber auch einfacher aufgebaut sein, das heißt es werden keine zusätzlichen Regulationssignale vor das Strukturgen insertiert und der natürli-

35

che Promotor mit seiner Regulation wird nicht entfernt. Statt dessen wird die natürliche Regulationssequenz so mutiert, dass keine Regulation mehr erfolgt und die Genexpression gesteigert oder verringert wird. Die Nukleinsäuresequenzen können in einer oder mehreren Kopien im Genkonstrukt enthalten sein.

5

Beispiele für brauchbare Promotoren sind: die Promotoren, *ddh*, *amy*, *lysC*, *dapA*, *lysA* aus *Corynebacterium glutamicum*, aber auch gram-positiven Promotoren *SPO2* wie sie in *Bacillus Subtilis* and Its Closest Relatives, Sonenshein, Abraham L., Hoch, James A., Losick, Richard; ASM Press, District of Columbia, Washington und Patek M. Eikmanns B.J. Patek J. Sahm H. Microbiology, 142 1297-309, 1996 beschrieben sind, oder aber auch *cos-*, *tac-*, *trp-*, *tet-*, *trp-tet-*, *lpp-*, *lac-*, *lpp-lac-*, *lacIq-*, *T7-*, *T5-*, *T3-*, *gal-*, *trc-*, *ara-*, *SP6-*, λ -PR- oder im λ -PL-Promotor, die vorteilhafterweise in gram-negativen Bakterien Anwendung finden. Bevorzugt ist auch die Verwendung induzierbarer Promotoren, wie z.B. licht- und insbesondere temperaturinduzierbarer Promotoren, wie der P_{P_1} -Promotor. Prinzipiell können alle natürlichen Promotoren mit ihren Regulationssequenzen verwendet werden. Darüber hinaus können auch synthetische Promotoren vorteilhaft verwendet werden.

10

15

Die genannten regulatorischen Sequenzen sollen die gezielte Expression der Nukleinsäuresequenzen ermöglichen. Dies kann beispielsweise je nach Wirtsorganismus bedeuten, dass das Gen erst nach Induktion exprimiert oder überexprimiert wird, oder dass es sofort exprimiert und/oder überexprimiert wird.

20

Die regulatorischen Sequenzen bzw. Faktoren können dabei vorzugsweise die Expression positiv beeinflussen und dadurch erhöhen oder erniedrigen. So kann eine Verstärkung der regulatorischen Elemente vorteilhafterweise auf der Transkriptionsebene erfolgen, indem starke Transkriptionssignale wie Promotoren und/oder "Enhancer" verwendet werden. Daneben ist aber auch eine Verstärkung der Translation möglich, indem beispielsweise die Stabilität der mRNA verbessert wird.

25

Die Herstellung einer Expressionskassette erfolgt durch Fusion eines geeigneten Promotors, einer geeigneten Shine-Dalgarno-Sequenz mit einer *metA*-Nukleotidsequenz sowie einem geeigneten Terminationssignal. Dazu verwendet man gängige Rekombinations- und Klonierungstechniken, wie sie beispielsweise in Current Protocols in Molecular Biology, 1993, John Wiley & Sons, Incorporated, New York New York, PCR Methods, Gelfand, David H., Innis, Michael A., Sninsky, John J. 1999, Academic Press, Incorporated, California, San Diego, ., PCR Cloning Protocols, Methods in Molecular Biology Ser., Vol. 192, 2nd ed., Humana Press, New

30

35

Jersey, Totowa. T. Maniatis, E.F. Fritsch und J. Sambrook, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989) sowie in T.J. Silhavy, M.L. Berman und L.W. Enquist, *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984) und in Ausubel, F.M. et al., *Current Protocols in Molecular Biology*, Greene Publishing Assoc. and Wiley Interscience (1987) beschrieben sind.

Das rekombinante Nukleinsäurekonstrukt bzw. Genkonstrukt wird zur Expression in einem geeigneten Wirtsorganismus vorteilhafterweise in einen wirtsspezifischen Vektor inseriert, der eine optimale Expression der Gene im Wirt ermöglicht. Vektoren sind dem Fachmann wohl bekannt und können beispielsweise aus "Cloning Vectors" (Pouwels P. H. et al., Hrsg, Elsevier, Amsterdam-New York-Oxford, 1985) entnommen werden. Unter Vektoren sind außer Plasmiden auch alle anderen dem Fachmann bekannten Vektoren, wie beispielsweise Phagen, Transposons, IS-Elemente, Phasmide, Cosmide, und lineare oder zirkuläre DNA zu verstehen. Diese Vektoren können autonom im Wirtsorganismus repliziert oder chromosomal repliziert werden.

Zur Verstärkung wurden erfindungsgemäße metA-Gene beispielhaft mit Hilfe von episomalen Plasmiden überexprimiert. Als Plasmide eignen sich solche, die in coryneformen Bakterien repliziert werden. Zahlreiche bekannte Plasmidvektoren, wie z. B. pZ1 (Menkel et al., *Applied and Environmental Microbiology* (1989) 64: 549-554), pEKEx1 (Eikmanns et al., *Gene* 102: 93-98 (1991)) oder pHS2-1 (Sonnen et al., *Gene* 107: 69-74 (1991)) beruhen auf den kryptischen Plasmiden pHM1519, pBL1 oder pGA1. Andere Plasmidvektoren, wie z. B. pCLiK5MCS, oder solche, die auf pCG4 (US-A 4,489,160) oder pNG2 (Serwold-Davis et al., *FEMS Microbiology Letters* 66, 119-124 (1990)) oder pAG1 (US-A 5,158,891) beruhen, können in gleicher Weise verwendet werden.

Weiterhin eignen sich auch solche Plasmidvektoren mit Hilfe derer man das Verfahren der Genamplifikation durch Integration in das Chromosom anwenden kann, so wie es beispielsweise von Remscheid et al. (*Applied and Environmental Microbiology* 60, 126-132 (1994)) zur Duplikation bzw. Amplifikation des *hom-thrB*-Operons beschrieben wurde. Bei dieser Methode wird das vollständige Gen in einen Plasmidvektor kloniert, der in einem Wirt (typischerweise *E. coli*), nicht aber in *C. glutamicum* replizieren kann. Als Vektoren kommen beispielsweise pSUP301 (Simon et al., *Bio/ Technology* 1, 784-791 (1983)), pK18mob oder pK19mob (Schäfer et al., *Gene* 145, 69-73 (1994)), Bernard et al., *Journal of Molecular Biology*, 234: 534-541 (1993)), pEM1 (Schrumpf et al. 1991, *Journal of Bacteriology* 173: 4510-4516) oder pBGS8 (Spratt et al., 1986, *Gene* 41: 337-342) in Frage. Der Plasmidvektor, der das zu amplifizierende Gen enthält, wird anschließend durch Transformation in den gewünschten Stamm von *C. glutamicum* überführt.

Methoden zur Transformation sind beispielsweise bei Thierbach et al. (Applied Microbiology and Biotechnology 29, 356-362 (1988)), Dunican und Shivnan (Biotechnology 7, 1067-1070 (1989)) und Tauch et al. (FEMS Microbiological Letters 123,343-347 (1994)) beschrieben.

- 5 Enzyme können durch Mutationen in den korrespondierenden Genen derart in ihrer Aktivität beeinflusst werden, dass es zu einer teilweisen oder vollständigen Verringerung der Reaktionsgeschwindigkeit der enzymatischen Reaktion kommt. Beispiele für solche Mutationen sind dem Fachmann bekannt (Motoyama H. Yano H. Terasaki Y. Anazawa H. Applied & Environmental Microbiology. 67:3064-70, 2001, Eikmanns BJ. Eggeling L. Sahm H. Antonie van Leeuwenhoek. 10 64:145-63, 1993-94.)

Zusätzlich kann es für die Produktion von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, vorteilhaft sein, neben einer Expression bzw. Verstärkung eines erfindungsgemäßen metA-Gen eines oder mehrere Enzyme des jeweiligen Biosyntheseweges, des Cystein-
15 Stoffwechselwegs, der Aspartatsemialdehyd-Synthese, der Glykolyse, der Anaplerotik, des Pentose-Phosphat-Stoffwechsels, des Zitronensäure-Zyklus oder des Aminosäure-Exports zu verstärken.

- So kann für die Herstellung von schwefelhaltige Feinchemikalien, insbesondere L-Methionin,
20 eines oder mehrere der folgenden Gene verstärkt sein:
- das für eine Aspartatkinase kodierende Gen lysC (EP 1 108 790 A2; DNA-SEQ NO. 281),
 - das für eine Aspartat-Semialdehyd Dehydrogenase kodierende Gen asd (EP 1 108 790 A2; DNA-SEQ NO. 282),
 - das für die Glycerinaldehyd-3-Phosphat Dehydrogenase kodierende Gen gap (Eikmanns
25 (1992), Journal of Bacteriology 174: 6076-6086),
 - das für die 3-Phosphoglycerat Kinase kodierende Gen pgk (Eikmanns (1992), Journal of Bacteriology 174: 6076-6086),
 - das für die Pyruvat Carboxylase kodierende Gen pyc (Eikmanns (1992), Journal of Bacteriology 174: 6076-6086),
 - 30 - das für die Triosephosphat Isomerase kodierende Gen tpi (Eikmanns (1992), Journal of Bacteriology 174: 6076-6086),
 - das für die Methionin Synthase kodierende Gen meth (EP 1 108 790 A2),
 - das für die Cystathionin-gamma-Synthase kodierende Gen metB (EP 1 108 790 A2; DNA-SEQ NO. 3491),
 - 35 - das für die Cystathionin-gamma-Lyase kodierende Gen metC (EP 1 108 790 A2; DNA-SEQ NO. 3061),

- das für die Serin-Hydroxymethyltransferase kodierende Gen glyA (EP 1 108 790 A2; DNA-SEQ NO. 1110),
- das für die O-Acetylhomoserin-Sulfhydrylase kodierende Gen metY (EP 1 108 790 A2; DNA-SEQ NO. 726),
- 5 - das für die Methylentetrahydrofolat-Reduktase kodierende Gen metF (EP 1 108 790 A2; DNA-SEQ NO. 2379),
- das für die Phosphoserin-Aminotransferase kodierende Gen serC (EP 1 108 790 A2; DNA-SEQ NO. 928)
- eines für die Phosphoserin-Phosphatase kodierende Gen serB (EP 1 108 790 A2; DNA-SEQ NO. 334, DNA-SEQ NO. 467, DNA-SEQ NO. 2767)
- 10 - das für die Serine Acetyl-Transferase kodierende Gen cysE (EP 1 108 790 A2; DNA-SEQ NO. 2818)
- das für eine Homoserin-Dehydrogenase kodierende Gen hom (EP 1 108 790 A2; DNA-SEQ NO. 1306)

15

So kann für die Herstellung von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, in coryneformen Bakterien, vorteilhaft sein, gleichzeitig wenigstens eines der nachfolgenden Gene zu mutieren, so dass die korrespondierenden Proteine, verglichen mit nicht mutierten Proteinen, in geringerem Maße oder nicht durch einen Stoffwechselmetaboliten in ihrer Aktivität beeinflusst werden:

20

- das für eine Aspartatkinase kodierende Gen lysC (EP 1 108 790 A2; DNA-SEQ NO. 281),
- das für die Pyruvat Carboxylase kodierende Gen pyc (Eikmanns (1992), Journal of Bacteriology 174: 6076-6086),
- 25 - das für die Methionin Synthase kodierende Gen metH (EP 1 108 790 A2),
- das für die Cystahionin-gamma-Synthase kodierende Gen metB (EP 1 108 790 A2; DNA-SEQ NO. 3491),
- das für die Cystahionin-gamma-Lyase kodierende Gen metC (EP 1 108 790 A2; DNA-SEQ NO. 3061),
- 30 - das für die Serin-Hydroxymethyltransferase kodierende Gen glyA (EP 1 108 790 A2; DNA-SEQ NO. 1110),
- das für die O-Acetylhomoserin-Sulfhydrylase kodierende Gen metY (EP 1 108 790 A2; DNA-SEQ NO. 726),
- das für die Methylentetrahydrofolat-Reduktase kodierende Gen metF (EP 1 108 790 A2; DNA-SEQ NO. 2379),
- 35 - das für die Phosphoserin-Aminotransferase kodierende Gen serC (EP 1 108 790 A2; DNA-

SEQ NO. 928)

- eines für die Phosphoserin-Phosphatase kodierende Gen serB (EP 1 108 790 A2; DNA-SEQ NO. 334, DNA-SEQ NO. 467, DNA-SEQ NO. 2767)

- das für die Serine Acetyl-Transferase kodierende Gen cysE (EP 1 108 790 A2; DNA-SEQ NO.

5 2818)

- das für eine Homoserin-Dehydrogenase kodierende Gen hom (EP 1 108 790 A2; DNA-SEQ NO. 1306)

10 Weiterhin kann es für die Produktion von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, vorteilhaft sein, zusätzlich zur Expression bzw. Verstärkung eines der erfindungsgemäßen metA-Gene eines oder mehrere der folgenden Gene abzuschwächen, insbesondere deren Expression zu verringern, oder auszuschalten:

- das für die Homoserine-Kinase kodierende Gen thrB (EP 1 108 790 A2; DNA-SEQ NO. 3453)

15 - das für die Threonin Dehydratase kodierende Gen ilvA (EP 1 108 790 A2; DNA-SEQ NO. 2328)

- das für die Threonin Synthase kodierende Gen thrC (EP 1 108 790 A2; DNA-SEQ NO. 3486)

- das für die Meso-Diaminopimelat D-Dehydrogenase kodierende Gen ddh (EP 1 108 790 A2; DNA-SEQ NO. 3494)

20 - das für die Phosphoenolpyruvat-Carboxykinase kodierende Gen pck (EP 1 108 790 A2; DNA-SEQ NO. 3157)

- das für die Glucose-6-Phosphat-6-Isomerase kodierende Gen pgi (EP 1 108 790 A2; DNA-SEQ NO. 950)

- das für die Pyruvat-Oxidase kodierende Gen poxB (EP 1 108 790 A2; DNA-SEQ NO. 2873)

25 - das für die Dihydrodipicolinat Synthase kodierende Gen dapA (EP 1 108 790 A2; DNA-SEQ NO. 3476)

- das für die Dihydrodipicolinat Reduktase kodierende Gen dapB (EP 1 108 790 A2; DNA-SEQ NO. 3477)

- das für die Diaminopicolinat Decarboxylase kodierende Gen lysA (EP 1 108 790 A2; DNA-SEQ

30 NO. 3451)

Weiterhin kann es für die Produktion von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, vorteilhaft sein, zusätzlich zur Expression bzw. Verstärkung eines der erfindungsgemäßen metA-Gene in Coryneformen Bakterien gleichzeitig wenigstens eines der folgenden Ge-
35 ne so zu mutieren, dass die enzymatische Aktivität des korrespondierenden Proteins teilweise oder vollständig verringert wird:

- das für die Homoserine-Kinase kodierende Gen thrB (EP 1 108 790 A2; DNA-SEQ NO. 3453)
- das für die Threonin Dehydratase kodierende Gen ilvA (EP 1 108 790 A2; DNA-SEQ NO. 2328)
- 5 - das für die Threonin Synthase kodierende Gen thrC (EP 1 108 790 A2; DNA-SEQ NO. 3486)
- das für die Meso-Diaminopimelat D-Dehydrogenase kodierende Gen ddh (EP 1 108 790 A2; DNA-SEQ NO. 3494)
- das für die Phosphoenolpyruvat-Carboxykinase kodierende Gen pck (EP 1 108 790 A2; DNA-SEQ NO. 3157)
- 10 - das für die Glucose-6-Phosphat-6-Isomerase kodierende Gen pgi (EP 1 108 790 A2; DNA-SEQ NO. 950)
- das für die Pyruvat-Oxidase kodierende Gen poxB (EP 1 108 790 A2; DNA-SEQ NO. 2873)
- das für die Dihydrodipicolinat Synthase kodierende Gen dapA (EP 1 108 790 A2; DNA-SEQ NO. 3476)
- 15 - das für die Dihydrodipicolinat Reduktase kodierende Gen dapB (EP 1 108 790 A2; DNA-SEQ NO. 3477)
- das für die Diaminopicolinat Decarboxylase kodierende Gen lysA (EP 1 108 790 A2; DNA-SEQ NO. 3451)
- 20 Weiterhin kann es für die Produktion von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, vorteilhaft sein, neben der Expression bzw. Verstärkung eines erfindungsgemäßen metA-Gens unerwünschte Nebenreaktionen auszuschalten (Nakayama: "Breeding of Amino Acid Producing Microorganisms", in: Overproduction of Microbial Products, Krumphanzl, Sikyta, Vanek (eds.), Academic Press, London, UK, 1982).
- 25 Die erfindungsgemäß hergestellten Mikroorganismen können kontinuierlich oder diskontinuierlich im batch- Verfahren (Satzkultivierung) oder im fed batch (Zulaufverfahren) oder repeated fed batch Verfahren (repetitives Zulaufverfahren) zur Produktion von schwefelhaltige Feinchemikalien, insbesondere L-Methionin, kultiviert werden. Eine Zusammenfassung über bekannte Kultivierungsmethoden ist im Lehrbuch von Chmiel (Bioprozeßtechnik 1. Einführung in die Bioverfahrenstechnik (Gustav Fischer Verlag, Stuttgart, 1991)) oder im Lehrbuch von Storhas (Bioreaktoren und periphere Einrichtungen (Vieweg Verlag, Braunschweig/Wiesbaden, 1994)) zu finden.
- 30 Das zu verwendende Kulturmedium hat in geeigneter Weise den Ansprüchen der jeweiligen Stämme zu genügen. Beschreibungen von Kulturmedien verschiedener Mikroorganismen sind
- 35

im Handbuch "Manual of Methods für General Bacteriology" der American Society für Bacteriology (Washington D. C., USA, 1981) enthalten.

5 Diese erfindungsgemäß einsetzbaren Medien umfassen gewöhnlich eine oder mehreren Kohlenstoffquellen, Stickstoffquellen, anorganische Salze, Vitamine und/oder Spurenelemente.

10 Bevorzugte Kohlenstoffquellen sind Zucker, wie Mono-, Di- oder Polysaccharide. Sehr gute Kohlenstoffquellen sind beispielsweise Glucose, Fructose, Mannose, Galactose, Ribose, Sorbose, Ribulose, Lactose, Maltose, Saccharose, Raffinose, Stärke oder Cellulose. Man kann Zucker auch über komplexe Verbindungen, wie Melassen, oder andere Nebenprodukte der Zucker-
15 Raffinierung zu den Medien geben. Es kann auch vorteilhaft sein, Gemische verschiedener Kohlenstoffquellen zuzugeben. Andere mögliche Kohlenstoffquellen sind Öle und Fette wie z. B. Sojaöl, Sonnenblumenöl, Erdnußöl und Kokosfett, Fettsäuren wie z. B. Palmitinsäure, Stearinsäure oder Linolsäure, Alkohole wie z. B. Glycerin, Methanol oder Ethanol und organische Säuren wie z. B. Essigsäure oder Milchsäure.

20 Stickstoffquellen sind gewöhnlich organische oder anorganische Stickstoffverbindungen oder Materialien, die diese Verbindungen enthalten. Beispielhafte Stickstoffquellen umfassen Ammoniak-Gas oder Ammoniumsalze, wie Ammoniumsulfat, Ammoniumchlorid, Ammoniumphosphat, Ammoniumcarbonat oder Ammoniumnitrat, Nitrate, Harnstoff, Aminosäuren oder komplexe Stickstoffquellen, wie Maisquellwasser, Sojamehl, Sojaprotein, Hefeextrakt, Fleischextrakt und andere. Die Stickstoffquellen können einzeln oder als Mischung verwendet werden.

25 Anorganische Salzverbindungen, die in den Medien enthalten sein können, umfassen die Chlorid-, Phosphor- oder Sulfatsalze von Calcium, Magnesium, Natrium, Kobalt, Molybdän, Kalium, Mangan, Zink, Kupfer und Eisen

30 Als Schwefelquelle für die Herstellung von schwefelhaltigen Feinchemikalien, insbesondere von Methionin, können anorganische schwefelhaltige Verbindungen wie beispielsweise Sulfate, Sulfite, Dithionite, Tetrathionate, Thiosulfate, Sulfide aber auch organische Schwefelverbindungen, wie Mercaptane und Thiole, verwendet werden.

Als Phosphorquelle können Phosphorsäure, Kaliumdihydrogenphosphat oder Dikaliumhydrogenphosphat oder die entsprechenden Natrium haltigen Salze verwendet werden.

35 Chelatbildner können zum Medium gegeben werden, um die Metallionen in Lösung zu halten.

Besonders geeignete Chelatbildner umfassen Dihydroxyphenole, wie Catechol oder Protocatechuat, oder organische Säuren, wie Citronensäure.

Die erfindungsgemäß eingesetzten Fermentationsmedien enthalten üblicherweise auch andere Wachstumsfaktoren, wie Vitamine oder Wachstumsförderer, zu denen beispielsweise Biotin, Riboflavin, Thiamin, Folsäure, Nikotinsäure, Panthothenat und Pyridoxin gehören. Wachstumsfaktoren und Salze stammen häufig von komplexen Medienkomponenten, wie Hefeextrakt, Melassen, Maisquellwasser und dergleichen. Dem Kulturmedium können überdies geeignete Vorstufen zugesetzt werden. Die genaue Zusammensetzung der Medienverbindungen hängt stark vom jeweiligen Experiment ab und wird für jeden spezifischen Fall individuell entschieden. Information über die Medienoptimierung ist erhältlich aus dem Lehrbuch "Applied Microbiol. Physiology, A Practical Approach" (Hrsg. P.M. Rhodes, P.F. Stanbury, IRL Press (1997) S. 53-73, ISBN 0 19 963577 3). Wachstumsmedien lassen sich auch von kommerziellen Anbietern beziehen, wie Standard 1 (Merck) oder BHI (Brain heart infusion, DIFCO) und dergleichen.

Sämtliche Medienkomponenten werden, entweder durch Hitze (20 min bei 1,5 bar und 121°C) oder durch Sterilfiltration, sterilisiert. Die Komponenten können entweder zusammen oder nötigenfalls getrennt sterilisiert werden. Sämtliche Medienkomponenten können zu Beginn der Anzucht zugegen sein oder wahlfrei kontinuierlich oder chargenweise hinzugegeben werden.

Die Temperatur der Kultur liegt normalerweise zwischen 15°C und 45°C, vorzugsweise bei 25°C bis 40°C und kann während des Experimentes konstant gehalten oder verändert werden. Der pH-Wert des Mediums sollte im Bereich von 5 bis 8,5, vorzugsweise um 7,0 liegen. Der pH-Wert für die Anzucht läßt sich während der Anzucht durch Zugabe von basischen Verbindungen wie Natriumhydroxid, Kaliumhydroxid, Ammoniak bzw. Ammoniakwasser oder sauren Verbindungen wie Phosphorsäure oder Schwefelsäure kontrollieren. Zur Kontrolle der Schaumentwicklung können Antischaummittel wie z. B. Fettsäurepolyglykolester, eingesetzt werden. Zur Aufrechterhaltung der Stabilität von Plasmiden können dem Medium geeignete selektiv wirkende Stoffe, wie z. B. Antibiotika, hinzugefügt werden. Um aerobe Bedingungen aufrechtzuerhalten, werden Sauerstoff oder Sauerstoff haltige Gasmischungen, wie z. B. Umgebungsluft, in die Kultur eingetragen. Die Temperatur der Kultur liegt normalerweise bei 20°C bis 45°C. Die Kultur wird solange fortgesetzt, bis sich ein Maximum des gewünschten Produktes gebildet hat. Dieses Ziel wird normalerweise innerhalb von 10 Stunden bis 160 Stunden erreicht.

Die so erhaltenen, insbesondere L-Methionin enthaltenden, Fermentationsbrühen haben üblicherweise eine Trockenmasse von 7,5 bis 25 Gew.-%.

Vorteilhaft ist außerdem auch, wenn die Fermentation zumindest am Ende, insbesondere jedoch über mindestens 30% der Fermentationsdauer zuckerlimitiert gefahren wird. Das heißt, dass während dieser Zeit die Konzentration an verwertbarem Zucker im Fermentationsmedium auf ≥ 0 bis 3 g/l gehalten, beziehungsweise abgesenkt wird.

Die Fermentationsbrühe wird anschließend weiterverarbeitet. Je nach Anforderung kann die Biomasse ganz oder teilweise durch Separationsmethoden, wie z. B. Zentrifugation, Filtration, Dekantieren oder einer Kombination dieser Methoden aus der Fermentationsbrühe entfernt oder vollständig in ihr belassen werden.

Anschließend kann die Fermentationsbrühe mit bekannten Methoden, wie z. B. mit Hilfe eines Rotationsverdampfers, Dünnschichtverdampfers, Fallfilmverdampfers, durch Umkehrosmose, oder durch Nanofiltration, eingedickt beziehungsweise aufkonzentriert werden. Diese aufkonzentrierte Fermentationsbrühe kann anschließend durch Gefriertrocknung, Sprühtrocknung, Sprühgranulation oder durch anderweitige Verfahren aufgearbeitet werden.

Es ist aber auch möglich die schwefelhaltigen Feinchemikalien, insbesondere L-Methionin, weiter aufzureinigen. Hierzu wird die produktthaltige Brühe nach dem Abtrennen der Biomasse einer Chromatographie mit einem geeigneten Harz unterworfen, wobei das gewünschte Produkt oder die Verunreinigungen ganz oder teilweise auf dem Chromatographieharz zurückgehalten werden. Diese Chromatographieschritte können nötigenfalls wiederholt werden, wobei die gleichen oder andere Chromatographieharze verwendet werden. Der Fachmann ist in der Auswahl der geeigneten Chromatographieharze und ihrer wirksamsten Anwendung bewandert. Das gereinigte Produkt kann durch Filtration oder Ultrafiltration konzentriert und bei einer Temperatur aufbewahrt werden, bei der die Stabilität des Produktes maximal ist.

Die Identität und Reinheit der isolierten Verbindung(en) kann durch Techniken des Standes der Technik bestimmt werden. Diese umfassen Hochleistungs-Flüssigkeitschromatographie (HPLC), spektroskopische Verfahren, Färbeverfahren, Dünnschichtchromatographie, NIRS, Enzymtest oder mikrobiologische Tests. Diese Analyseverfahren sind zusammengefaßt in: Patek et al. (1994) Appl. Environ. Microbiol. 60:133-140; Malakhova et al. (1996) Biotekhnologiya 11 27-32; und Schmidt et al. (1998) Bioprocess Engineer. 19:67-70. Ulmann's Encyclopedia of Industrial Chemistry (1996) Bd. A27, VCH: Weinheim, S. 89-90, S. 521-540, S. 540-547, S. 559-566, 575-581 und S. 581-587; Michal, G (1999) Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, John Wiley and Sons; Fallon, A. et al. (1987) Applications of HPLC in Bioche-

mistry in: Laboratory Techniques in Biochemistry and Molecular Biology, Bd. 17.

Die Erfindung wird nun anhand der folgenden nicht-limitierenden Beispiele und unter Bezugnahme auf beiliegende Figuren näher beschrieben. Dabei zeigt

5

Figur 1 die Plasmidkarte zu Plasmid pClysC;

Figur 2 die Plasmidkarte zu Plasmid pCISlysCthr311le;

Figur 3 die Plasmidkarte zu Plasmid pC_metaA_Cd.

10 Restriktionsschnittstellen mit der entsprechenden Positionsangabe in Klammern sind in den Plasmidkarten angegeben. Wesentliche Sequenzabschnitte sind fettgedruckt beschrieben. KanR steht für Kanamycin-Resistenzgen; ask steht für Aspartatkinasegen.

Beispiel 1: Konstruktion von pCLiK5MCS

15

Zunächst wurden Ampicillinresistenz und Replikationsursprung des Vektors pBR322 mit den Oligonukleotiden p1.3 (SEQ ID NO:47) und p2.3 (SEQ ID NO:48) mit Hilfe der Polymerase-Kettenreaktion (PCR) amplifiziert.

20

p1.3 (SEQ ID NO:47)

5'-CCCGGGATCCGCTAGCGGCGCGCCGGCCGGCCCGGTGTGAAATACCGCACAG-3'

p2.3 (SEQ ID NO:48)

5'-TCTAGACTCGAGCGGCCGCGGCCGGCCTTTAAATTGAAGACGAAAGGGCCTCG-3'

25

Neben den zu pBR322 komplementären Sequenzen, enthält das Oligonukleotid p1.3 (SEQ ID NO:47) in 5'-3' Richtung die Schnittstellen für die Restriktionsendonukleasen SmaI, BamHI, NheI und AscI und das Oligonukleotid p2.3 (SEQ ID NO:48) in 5'-3' Richtung die Schnittstellen für die Restriktionsendonukleasen XbaI, XhoI, NotI und DraI. Die PCR Reaktion wurde nach Standardmethode wie Innis et al. (PCR Protocols. A Guide to Methods and Applications, Academic Press (1990)) mit PfuTurbo Polymerase (Stratagene, La Jolla, USA) durchgeführt. Das erhaltene DNA Fragment mit einer Größe von ungefähr 2,1 kb wurde mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt. Die stumpfen Enden des DNA-Fragmentes wurden mit dem Rapid DNA Ligation Kit (Roche

30

35

Diagnosics, Mannheim) nach Angaben des Herstellers miteinander ligiert und der Ligationsan-

satz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben(1989)), in kompetente E.coli XL-1Blue (Stratagene, La Jolla, USA) transformiert. Eine Selektion auf Plasmid tragende Zellen wurde durch das Ausplattieren auf Ampicillin (50µg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.

5

Die Plasmid-DNA eines individuellen Klons wurde mit dem Qiaprep Spin Miniprep Kit (Qiagen, Hilden) nach Angaben des Herstellers isoliert und über Restriktionsverdaus überprüft. Das so erhaltene Plasmid erhält den Namen pCLiK1.

- 10 Ausgehend vom Plasmid pWLT1 (Liebl et al., 1992) als Template für eine PCR Reaktion wurde mit den Oligonukleotiden neo1 (SEQ ID NO:49) und neo2 (SEQ ID NO:50) eine Kanamycin-Resistenzcassette amplifiziert.

neo1 (SEQ ID NO:49):

- 15 5'-GAGATCTAGACCCGGGGATCCGCTAGCGGGCTGCTAAAGGAAGCGGA-3'

neo2 (SEQ ID NO:50):

5'-GAGAGGCGCGCCGCTAGCGTGGGCGAAGAACTCCAGCA-3'

- 20 Neben den zu pWLT1 komplementären Sequenzen, enthält das Oligonukleotid neo1 in 5'-3' Richtung die Schnittstellen für die Restriktionsendonukleasen XbaI, SmaI, BamHI, NheI und das Oligonukleotid neo2 (SEQ ID NO:50) in 5'-3' Richtung die Schnittstellen für die Restriktionsendonukleasen AscI und NheI. Die PCR Reaktion wurde nach Standardmethode wie Innis et al. (PCR Protocols. A Guide to Methods and Applications, Academic Press (1990)) mit PfuTurbo
- 25 Polymerase (Stratagene, La Jolla, USA) durchgeführt. Das erhaltene DNA Fragment mit einer Größe von ungefähr 1,3 kb wurde mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt. Das DNA-Fragment wurde mit den Restriktionsendonukleasen XbaI und AscI (New England Biolabs, Beverly, USA) geschnitten und im Anschluß daran erneut mit dem GFX™PCR, DNA and Gel Band Purification
- 30 Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt. Der Vektor pCLiK1 wurde ebenfalls mit den Restriktionsendonukleasen XbaI und AscI geschnitten und mit alkalischer Phosphatase (Roche Diagnostics, Mannheim) nach Angaben des Herstellers dephosphoryliert. Nach Elektrophorese in einem 0,8%igen Agarosegel wurde der linearisierte Vektor (ca. 2,1kb) mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharma-
- 35 cia, Freiburg) nach Angaben des Herstellers isoliert. Dieses Vektor-Fragment wurde mit Hilfe des Rapid DNA Ligation Kit (Roche Diagnostics, Mannheim) nach Angaben des Herstellers mit

dem geschnittenen PCR Fragment ligiert und der Ligationsansatz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben(1989)), in kompetente E.coli XL-1Blue (Stratagene, La Jolla, USA) transformiert. Eine Selektion auf Plasmid tragende Zellen wurde durch das Ausplattieren auf Ampicillin (50µg/ml) und Kanamycin (20µg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.

Die Plasmid-DNA eines individuellen Klons wurde mit dem Qiaprep Spin Miniprep Kit (Qiagen, Hilden) nach Angaben des Herstellers isoliert und über Restriktionsverdaus überprüft. Das so erhaltene Plasmid erhält den Namen pCLiK2.

Der Vektor pCLiK2 wurde mit der Restriktionsendonuklease DraI (New England Biolabs, Beverly, USA) geschnitten. Nach Elektrophorese in einem 0,8%igen Agarosegel wurde ein ca. 2,3 kb großes Vektorfragment mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers isoliert. Dieses Vektor-Fragment wurde mit Hilfe des Rapid DNA Ligation Kit (Roche Diagnostics, Mannheim) nach Angaben des Herstellers religiert und der Ligationsansatz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben (1989)), in kompetente E.coli XL-1Blue (Stratagene, La Jolla, USA) transformiert. Eine Selektion auf Plasmid tragende Zellen wurde durch das Ausplattieren auf Kanamycin (20µg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.

Die Plasmid-DNA eines individuellen Klons wurde mit dem Qiaprep Spin Miniprep Kit (Qiagen, Hilden) nach Angaben des Herstellers isoliert und über Restriktionsverdaus überprüft. Das so erhaltene Plasmid erhält den Namen pCLiK3.

Ausgehend vom Plasmid pWLQ2 (Liebl et al., 1992) als Template für eine PCR Reaktion wurde mit den Oligonukleotiden cg1 ((SEQ ID NO:51) und cg2 (SEQ ID NO:52) der Replikationsursprung pHM1519 amplifiziert.

cg1 (SEQ ID NO:51):
5'-GAGAGGGCGGCCGCGCAAAGTCCCGCTTCGTGAA-3'

cg2 (SEQ ID NO:52):
5'-GAGAGGGCGGCCGCTCAAGTCGGTCAAGCCACGC-3'

Neben den zu pWLQ2 komplementären Sequenzen, enthalten die Oligonukleotide cg1 (SEQ ID

- NO:51) und cg2 (SEQ ID NO:52) Schnittstellen für die Restriktionsendonuklease NotI. Die PCR Reaktion wurde nach Standardmethode wie Innis et al. (PCR Protocols. A Guide to Methods and Applications, Academic Press (1990)) mit PfuTurbo Polymerase (Stratagene, La Jolla, USA) durchgeführt. Das erhaltene DNA Fragment mit einer Größe von ungefähr 2,7 kb wurde mit dem
- 5 GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt. Das DNA-Fragment wurde mit der Restriktionsendonuklease NotI (New England Biolabs, Beverly, USA) geschnitten und im Anschluß daran erneut mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt. Der Vektor pCLiK3 wurde ebenfalls mit der Restriktionsendonuklease
- 10 NotI geschnitten und mit alkalischer Phosphatase (Roche Diagnostics, Mannheim)) nach Angaben des Herstellers dephosphoryliert. Nach Elektrophorese in einem 0,8%igen Agarosegel wurde der linearisierte Vektor (ca. 2,3kb) mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers isoliert. Dieses Vektor-Fragment wurde mit Hilfe des Rapid DNA Ligation Kit (Roche Diagnostics, Mannheim) nach An-
- 15 gaben des Herstellers mit dem geschnittenen PCR Fragment ligiert und der Ligationsansatz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben(1989)), in kompetente E.coli XL-1Blue (Stratagene, La Jolla, USA) transformiert. Eine Selektion auf Plasmid tragende Zellen wurde durch das Ausplattieren auf Kanamycin (20µg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.
- 20 Die Plasmid-DNA eines individuellen Klons wurde mit dem Qiaprep Spin Miniprep Kit (Qiagen, Hilden) nach Angaben des Herstellers isoliert und über Restriktionsverdaus überprüft. Das so erhaltene Plasmid erhält den Namen pCLiK5.
- 25 Für die Erweiterung von pCLiK5 um eine „multiple cloning site“ (MCS) wurden die beide synthetischen, weitestgehend komplementären Oligonukleotide HS445 ((SEQ ID NO:53) und HS446 (SEQ ID NO:54), die Schnittstellen für die Restriktionsendonukleasen SwaI, XhoI, AatI, ApaI, Asp718, MluI, NdeI, SpeI, EcoRV, Sall, ClaI, BamHI, XbaI und SmaI enthalten, durch gemeinsames erhitzen auf 95°C und langsames abkühlen zu einem doppelsträngigen DNA-Fragment
- 30 vereinigt.

HS445 (SEQ ID NO:53):

5'-TCGAATTTAAATCTCGAGAGGCCTGACGTCGGGCCCCGGTACCACGCGTCATATGACTAG
TTCGGACCTAGGGATATCGTCGACATCGATGCTCTTCTGCGTTAATTAACAATTGGGATCC
35 TCTAGACCCGGGATTAAAT-3'

HS446 (SEQ ID NO:54):

5'-GATCATTAAATCCCGGGTCTAGAGGATCCCAATTGTTAATTAACGCAGAAGAGCATCGA
TGTCGACGATATCCCTAGGTCCGAAGTAGTCATATGACGCGTGGTACCGGGCCCGACGTC
AGGCCTCTCGAGATTAAAT-3'

5

Der Vektor pCLiK5 wurde mit den Restriktionsendonuklease XhoI und BamHI (New England Biolabs, Beverly, USA) geschnitten und mit alkalischer Phosphatase (I (Roche Diagnostics, Mannheim)) nach Angaben des Herstellers dephosphoryliert. Nach Elektrophorese in einem 0,8%igen Agarosegel wurde der linearisierte Vektor (ca. 5,0 kb) mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers isoliert. Dieses Vektor-Fragment wurde mit Hilfe des Rapid DNA Ligation Kit (Roche Diagnostics, Mannheim) nach Angaben des Herstellers mit dem synthetischen Doppelsträngigen DNA-Fragment ligiert und der Ligationsansatz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben(1989)), in kompetente E.coli XL-1Blue (Stratagene, La Jolla, USA) transformiert. Eine Selektion auf Plasmid tragende Zellen wurde durch das Ausplattieren auf Kanamycin (20µg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.

10

15

20

Die Plasmid-DNA eines individuellen Klons wurde mit dem Qiaprep Spin Miniprep Kit (Qiagen, Hilden) nach Angaben des Herstellers isoliert und über Restriktionsverdaus überprüft. Das so erhaltene Plasmid erhält den Namen pCLiK5MCS.

25

Sequenzierungsreaktionen wurden nach Sanger et al. (1977) Proceedings of the National Academy of Sciences USA 74:5463-5467 durchgeführt. Die Sequenzierreaktionen wurden mittels ABI Prism 377 (PE Applied Biosystems, Weiterstadt) aufgetrennt und ausgewertet.

Das entstandene Plasmid pCLiK5MCS ist als SEQ ID NO: 57 aufgeführt.

Beispiel 2: Konstruktion von pCLiK5MCS integrativ sacB

30

Ausgehend vom Plasmid pK19mob (Schäfer et al., Gene 145,69-73(1994)) als Template für eine PCR Reaktion wurde mit den Oligonukleotiden BK1732 und BK1733 das Bacillus subtilis sacB Gen (kodierend für Levan Sucrase) amplifiziert.

35

BK1732 (SEQ ID NO:55):

5'-GAGAGCGGCCGCGGATCCTTTTAAACCATCAC-3'

BK1733 (SEQ ID NO:56):

5'-AGGAGCGGCCGCCATCGGCATTTTCTTTTGCG-3'

5 Neben den zu pEK19mobsac komplementären Sequenzen, enthalten die Oligonukleotide BK1732 und BK1733 Schnittstellen für die Restriktionsendonuklease NotI. Die PCR Reaktion wurde nach Standardmethode wie Innis et al. (PCR Protocols. A Guide to Methods and Applications, Academic Press (1990)) mit PfuTurbo Polymerase (Stratagene, La Jolla, USA) durchgeführt. Das erhaltene DNA Fragment mit einer Größe von ungefähr 1,9 kb wurde mit dem
10 GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt. Das DNA-Fragment wurde mit der Restriktionsendonuklease NotI (New England Biolabs, Beverly, USA) geschnitten und im Anschluß daran erneut mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt.

15 Der Vektor pCLiK5MCS (hergestellt gemäß Beispiel 1) wurde ebenfalls mit der Restriktionsendonuklease NotI geschnitten und mit alkalischer Phosphatase I (Roche Diagnostics, Mannheim) nach Angaben des Herstellers dephosphoryliert. Nach Elektrophorese in einem 0,8%igen Agarosegel wurde ein ungefähr 2,4 kb großes Vektorfragment mit dem GFX™PCR, DNA and
20 Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers isoliert. Dieses Vektor-Fragment wurde mit Hilfe des Rapid DNA Ligation Kit (Roche Diagnostics, Mannheim) nach Angaben des Herstellers mit dem geschnittenen PCR Fragment ligiert und der Ligationsansatz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben(1989)), in kompetente E.coli XL-1Blue (Stratagene,
25 La Jolla, USA) transformiert. Eine Selektion auf Plasmid tragende Zellen wurde durch das Ausplattieren auf Kanamycin (20µg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.

Die Plasmid-DNA eines individuellen Klons wurde mit dem Qiaprep Spin Miniprep Kit (Qiagen, Hilden) nach Angaben des Herstellers isoliert und über Restriktionsverdaus überprüft. Das so
30 erhaltene Plasmid erhält den Namen pCLiK5MCS integrativ sacB.

Sequenzierungsreaktionen wurden nach Sanger et al. (1977) Proceedings of the National Academy of Sciences USA 74:5463-5467 durchgeführt. Die Sequenzierreaktionen wurden mittels ABI Prism 377 (PE Applied Biosystems, Weiterstadt) aufgetrennt und ausgewertet.

35

Das entstandene Plasmid pCLiK5MCS integrativ sacB ist als SEQ ID NO: 58 aufgeführt.

Weitere Vektoren die zur erfindungsgemäßen Expression oder Überproduktion von metA-Genen geeignet sind, können in analoger Weise hergestellt werden.

5 **Beispiel 3: Isolierung des lysC Gens aus dem C. glutamicum Stamm LU1479**

Im ersten Schritt der Stammkonstruktion soll ein allelischer Austausch des lysC Wildtypgens, kodierend für das Enzym Aspartatkinase, in C. glutamicum ATCC13032, im folgenden LU1479 genannt, durchgeführt werden. Dabei soll im LysC Gen ein Nukleotidaustausch durchgeführt werden, so dass im resultierenden Protein die Aminosäure Thr an der Position 311 durch die Aminosäure Ile ausgetauscht ist.

Ausgehend von der chromosomalen DNA aus LU1479 als Template für eine PCR Reaktion wurde mit den Oligonukleotidprimern SEQ ID NO:59 und SEQ ID NO:60 lysC mit Hilfe des Pfu-Turbo PCR Systems (Stratagene USA) nach Angaben des Herstellers amplifiziert. Chromosomale DNA aus C. glutamicum ATCC 13032 wurde nach Tauch et al. (1995) Plasmid 33:168-179 oder Eikmanns et al. (1994) Microbiology 140:1817-1828 präpariert. Das amplifizierte Fragment wird an seinem 5'-Ende von einem Sall Restriktionsschnitt und an seinem 3'-Ende von einem MluI Restriktionsschnitt flankiert. Vor der Klonierung wurde das amplifizierte Fragment durch diese beiden Restriktionsenzyme verdaut und mit GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) aufgereinigt.

SEQ ID NO:59

5'-GAGAGAGAGACGCGTCCCAGTGGCTGAGACGCATC -3'

25

SEQ ID NO:60

5'-CTCTCTCTGTCTGACGAATTCAATCTTACGGCCTG-3'

Das erhaltenen Polynukleotid wurde über die Sall und MluI Restriktionsschnitte in pCLIK5 MCS integrativ SacB (im folgenden pCIS genannt; SEQ ID NO: 58 aus Beispiel 2) kloniert und in E.coli XL-1 blue transformiert. Eine Selektion auf Plasmid-tragende Zellen wurde durch das Ausplattieren auf Kanamycin (20µg/ml)-haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht. Das Plasmid wurden isoliert und durch Sequenzierung die erwartete Nukleotidsequenz bestätigt. Die Präparation der Plasmid-DNA wurde nach Methoden und mit Materialien der Firma Quiagen durchgeführt. Sequenzierungsreaktionen wurden nach Sanger et al. (1977) Proceedings of the National Academy of Sciences USA 74:5463-5467 durchgeführt. Die Sequenzier-

reaktionen wurden mittels ABI Prism 377 (PE Applied Biosystems, Weiterstadt) aufgetrennt und ausgewertet. Das erhaltene Plasmid pCIS lysC ist als SEQ ID NO:61 aufgeführt. Die entsprechende Plasmidkarte ist in Figur 1 dargestellt.

5 Die Sequenz SEQ ID NO:61 umfasst die folgenden wesentlichen Teilbereiche:

LOCUS pCIS\lysC 5860 bp DNA circular

FEATURES Location/Qualifiers

CDS¹⁾ 155..1420

/vntifkey="4"

/label=lysC

10 CDS complement²⁾(3935..5356)

/vntifkey="4"

/label=sacB\(*Bacillus subtilis*)

15 promoter complement(5357..5819)

/vntifkey="30"

/label=Promotor\sacB

C_region complement(3913..3934)

/vntifkey="2"

/label=sacB\downstreambereich

20 CDS 1974..2765

/vntifkey="4"

/label=Kan\R

CDS complement(3032..3892)

/vntifkey="4"

25 /label=Ori\-EC\(*pMB*)

¹⁾ kodierende Sequenz

²⁾ auf Komplementärstrang

Beispiel 4: Mutagenese des lysC Gens aus *C. glutamicum*

30 Die gerichtete Mutagenese des lysC Gens aus *C. glutamicum* (Beispiel 3) wurde mit dem QuickChange Kit (Fa. Stratagene/USA) nach Angaben des Herstellers durchgeführt. Die Mutagenese wurde im Plasmid pCIS lysC, SEQ ID NO:61 durchgeführt. Für den Austausch von thr311 nach 311ile mit Hilfe der Quickchange Methode (Stratagene) wurden folgende Oligonukleotidprimer synthetisiert

35

SEQ ID NO:62

5'-CGGCACCACCGACATCATCTTCACCTGCCCTCGTTCCG -3'

SEQ ID NO:63

5'-CGGAACGAGGGCAGGTGAAGATGATGTCGGTGGTGCCG -3'

- 5 Der Einsatz dieser Oligonukleotidprimer in der Quickchange Reaktion führt in dem lysC Gen zu einem Austausch des Nukleotids in Position 932 (von C nach T) (vgl. SEQ ID NO:64) und im korrespondierenden Enzym zu einem Aminosäuresubstitution in Position 311 (Thr→Ile) (vgl. SEQ ID NO:65). Der resultierende Aminosäureaustausch Thr311Ile im lysC Gen wurde nach Transformation in E.coli XL1-blue und Plasmidpräparation durch Sequenzierung bestätigt. Das
- 10 Plasmid erhielt die Bezeichnung pCIS lysC thr311ile und ist als SEQ ID NO:66 aufgeführt. Die entsprechende Plasmidkarte ist in Figur 2 dargestellt.

Die Sequenz SEQ ID NO:66 umfasst die folgenden wesentlichen Teilbereiche:

15 LOCUS pCIS\lysC\thr311ile 5860 bp DNA circular
 FEATURES Location/Qualifiers
 CDS¹⁾ 155..1420
 /vntifkey="4"
 /label=lysC
 20 CDS complement²⁾(3935..5356)
 /vntifkey="4"
 /label=sacB\Bacillus\subtilis)
 promoter complement(5357..5819)
 25 /vntifkey="30"
 /label=Promotor\sacB
 C_region complement(3913..3934)
 /vntifkey="2"
 /label=sacB\downstreambereich
 30 CDS 1974..2765
 /vntifkey="4"
 /label=Kan\R
 CDS complement(3032..3892)
 /vntifkey="4"
 35 /label=Ori\EC\pMB)

¹⁾ kodierende Sequenz

²⁾ auf Komplementärstrang

Das Plasmid pCIS lysC thr311ile wurde in *C. glutamicum* LU1479 mittels Elektroporation wie bei Liebl, et al. (1989) FEMS Microbiology Letters 53:299-303 beschrieben, transformiert. Modifikationen des Protokolls sind in DE-A-10046870 beschrieben. Die chromosomale Anordnung des lysC-Lokus einzelner Transformanten wurde mit Standardmethoden durch Southernblot und Hybridisierung, wie in Sambrook et al. (1989), Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben, überprüft. Dadurch wurde sichergestellt, dass es sich bei den Transformanten um solche handelt, die das transformierte Plasmid durch homologe Rekombination am lysC-Lokus integriert haben. Nach Wachstum solcher Kolonien über Nacht in Medien, die kein Antibiotikum enthielten, wurden die Zellen auf ein Saccharose-CM-Agarmedium (10% Saccharose) ausplattiert und bei 30°C für 24 Stunden inkubiert.

Da das im Vektor pCIS lysC thr311ile enthaltende sacB Gen Saccharose in ein toxisches Produkt umwandelt, können nur solche Kolonien anwachsen, die das sacB Gen durch einen zweiten homologen Rekombinationsschritt zwischen dem Wildtyp lysC Gen und dem mutierten Gen lysC thr311ile deletiert haben. Während der homologen Rekombination kann entweder das Wildtyp Gen oder das mutierte Gen zusammen mit dem sacB Gen deletiert werden. Wenn das sacB Gen zusammen mit dem Wildtyp Gen entfernt wird, resultiert eine mutierte Transformante.

Anwachsende Kolonien wurden gepickt, und auf eine Kanamycin-sensitiven Phänotyp hin untersucht. Klone mit deletiertem SacB Gen müssen gleichzeitig Kanamycin-sensitives Wachstumsverhalten zeigen. Solche Kan-sensitiven Klone wurde im einem Schüttelkolben auf ihre Lysin-Produktivität hin untersucht (siehe Beispiel 6). Zum Vergleich wurde der nichtbehandelte Stamm LU1479 angezogen. Klone mit einer gegenüber der Kontrolle erhöhten Lysin-Produktion wurden selektiert, chromosomale DNA wurde gewonnen und der entsprechende Bereich des lysC Gens wurde durch eine PCR-Reaktion amplifiziert und sequenziert. Ein solcher Klon mit der Eigenschaft erhöhter Lysin-Synthese und nachgewiesener Mutation in lysC an der Stelle 932 wurde mit LU1479 lysC 311ile bezeichnet).

Beispiel 5: Herstellung Ethionin-resistenter *C. glutamicum* Stämme

Im zweiten Schritt der Stammkonstruktion wurde der erhaltene Stamm LU1479 lysC 311ile (Beispiel 4) behandelt, um eine Ethionin-Resistenz (Kase, H. Nakayama K.Agr. Biol. Chem. 39 153-106 1975 L-methionine production by methionine analog-resistant mutants of *Corynebacterium glutamicum*) zu induzieren: Eine Übernachtskultur in BHI-Medium (Difco) wurde in Citratpuffer (50mM pH 5,5) gewaschen und bei 30°C für 20 min mit N-Methyl-nitrosoguanidin (10mg/ml in 50mM Citrat pH5,5) behandelt. Nach der Behandlung mit dem chemischen Mutagen N-Methyl-

nitrosoguanidin wurden die Zellen gewaschen (Citratpuffer 50mM pH 5,5) und auf ein Medium plattiert, das aus folgenden Komponenten, berechnet auf 500ml, zusammengesetzt war: 10g $(\text{NH}_4)_2\text{SO}_4$, 0.5g KH_2PO_4 , 0.5g K_2HPO_4 , 0.125g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 21g MOPS, 50mg CaCl_2 , 15mg Proteokatechuat, 0,5mg Biotin, 1mg Thiamin, 5g/l D,L-Ethionin (Sigma Chemicals Deutschland), pH 7,0. Außerdem enthielt das Medium 0.5ml einer Spurensalzlösung aus: 10g/l $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, 1g/l $\text{MnSO}_4 \cdot \text{H}_2\text{O}$, 0.1g/l $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$, 0.02g/l CuSO_4 , 0.002g/l $\text{NiCl}_2 \cdot 6\text{H}_2\text{O}$. Alle Salze wurden in 0,1M HCl gelöst. Das fertig zusammengestellte Medium wurde sterilfiltriert und nach Zugabe von 40ml steriler 50% Glucoselösung, mit flüssigem sterilem Agar in einer Endkonzentration von 1,5% Agar versetzt und in Kulturschalen ausgegossen.

10

Auf Platten mit dem beschriebenen Medium wurden mutagenisierte Zellen aufgebracht und 3-7 Tage bei 30°C inkubiert. Erhaltene Klone wurden isoliert, mindestens einmal auf dem Selektionsmedium vereinzelt und dann auf ihre Methionin-Produktivität in einem Schüttelkolben in Medium II untersucht (siehe Beispiel 6)

15

Beispiel 6: Herstellung von Methionin mit dem Stamm LU1479 lysC 311ile ET-16.

Die in Beispiel 5 hergestellten Stämme wurden auf einer Agar-Platte mit CM-Medium für 2 Tag bei 30°C angezogen.

20

CM-Agar:

10,0 g/l D-Glucose, 2,5 g/l NaCl, 2,0 g/l Harnstoff, 10,0 g/l Bacto Pepton (Difco), 5,0 g/l Yeast Extract (Difco), 5,0 g/l Beef Extract (Difco), 22,0 g/l Agar (Difco), autoklaviert (20 min., 121°C)

25

Anschließend wurden die Zellen von der Platte abgekratzt und in Saline resuspendiert. Für die Hauptkultur wurden 10 ml Medium II und 0,5 g autoklaviertes CaCO_3 (Riedel de Haen) in einem 100 ml Erlenmeyerkolben mit der Zellsuspension bis zu einer OD600nm von 1,5 beimpft und für 72h auf einem Orbitalschüttler mit 200 Upm bei 30°C inkubiert.

30

Medium II:

40g/l Saccharose

60g/l Melasse (auf 100% Zuckergehalt berechnet)

10g/l $(\text{NH}_4)_2\text{SO}_4$

0.4g/l $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$

35

0.6g/l KH_2PO_4

0.3mg/l Thiamin*HCl
1mg/l Biotin (aus einer 1 mg/ml steril filtrierten Stammlösung die mit NH₄OH auf pH 8,0 eingestellt wurde)
2mg/l FeSO₄
5 2mg/l MnSO₄
mit NH₄OH auf pH 7,8 eingestellt, autoklaviert (121°C, 20 min). Zusätzlich wird Vitamin B12 (Hydroxycobalamin Sigma Chemicals) aus einer Stammlösung (200 µg/ml, steril filtriert) bis zu einer Endkonzentration von 100 µg/l zugegeben

10 Gebildetes Methionin, sowie andere Aminosäuren in der Kulturbrühe wurde mit Hilfe der Aminosäuresäure-Bestimmungsmethode von Agilent auf einer Agilent 1100 Series LC System HPLC. Eine Derivatisierung vor der Säulentrennung mit Ortho-Phthalaldehyd erlaubte die Quantifizierung der gebildeten Aminosäuren. Die Auftrennung des Aminosäuregemisch fand auf einer Hypersil AA-Säule (Agilent) statt.

15 Solche Klone wurden isoliert, deren Methionin-Produktivität mindestens doppelt so hoch war, wie die des Ausgangsstamm LU1479 lysC 311ile. Ein solcher Klon wurde für die weiteren Versuche eingesetzt und bekam die Bezeichnung LU1479 lysC 311ile ET-16.

20 **Beispiel 7:** Klonierung von metA aus *Corynebacterium diphtheriae* und Klonierung in das Plasmid pC metA_Cd

Chromosomale DNA von *Corynebacterium diphtheriae* wurde von der American Type Strain Culture Collection (ATCC, Atlanta-USA) mit der Bestellnummer 700971D aus dem Stamm ATCC
25 700971 bezogen.

Mit den Oligonukleotidprimer SEQ ID NO: 67 und SEQ ID NO:68, der chromosomalen DNA aus *C. diphtheriae* als Template und Pfu Turbo Polymerase (Fa. Stratagene) wurde mit Hilfe der Polymerase-Kettenreaktion (PCR) nach Standardmethoden wie Innis et al. (1990) PCR Protocols. A Guide to Methods and Applications, Academic Press ein DNA Fragment von ca. 1,4 kb amplifiziert, welches das metA Gen inklusive eines nichtkodierenden 5'-Bereiches (Promotorregion) enthält. Das amplifizierte Fragment ist an seinem 5'-Ende von einer XhoI-Restriktionsschnittstelle und am 3'-Ende von einer NdeI-Restriktionsschnittstelle flankiert, welche über die Oligonukleotidprimer eingeführt wurden.

35

SEQ ID NO:67

5'-GAGACTCGAGGTTGGCTGGTCATCATAGG -3'

und

SEQ ID NO:68

5'-GAAGAGAGCATATGTCAGCGCTCTAGTTTGGTTC -3'

5

Das erhaltene DNA Fragment wurde mit dem GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) nach Angaben des Herstellers gereinigt. Im Anschluß daran wurde es mit den Restriktionsenzymen XhoI und NdeI (Roche Diagnostics, Mannheim) gespalten und gelelektrophoretisch aufgetrennt. Anschließend wurde das ca. 1,4 kb große DNA Fragment mit GFX™PCR, DNA and Gel Band Purification Kit (Amersham Pharmacia, Freiburg) aus der Agarose aufgereinigt.

10

Der Vektor pClik5MCS SEQ ID NO: 57, im folgenden pC genannt, wurde mit den Restriktionsenzymen XhoI und NdeI (Roche Diagnostics, Mannheim) geschnitten und ein ca. 5 kb großes Fragment nach elektrophoretischer Auftrennung mit GFX™PCR, DNA and Gel Band Purification Kit isoliert.

15

Das Vektorfragment wurde zusammen mit dem PCR-Fragment mit Hilfe des Rapid DNA Ligation Kit (Roche Diagnostics, Mannheim) nach Angaben des Herstellers ligiert und der Ligationssatz nach Standardmethoden wie in Sambrook et al. (Molecular Cloning. A Laboratory Manual, Cold Spring Harbor, beschrieben(1989)), in kompetente E.coli XL-1Blue (Stratagene, La Jolla, USA) transformiert. Eine Selektion auf Plasmid tragende Zellen wurde durch das Ausplattieren auf Kanamycin (20µg/ml) haltigen LB Agar (Lennox, 1955, Virology, 1:190) erreicht.

20

Die Präparation der Plasmid DNA wurde nach Methoden und mit Materialien der Fa. Quiagen durchgeführt. Sequenzierungsreaktionen wurden nach Sanger et al. (1977) Proceedings of the National Academy of Sciences USA 74:5463-5467 durchgeführt. Die Sequenzierreaktionen wurden mittels ABI Prism 377 (PE Applied Biosystems, Weiterstadt) aufgetrennt und ausgewertet.

25

Das entstandene Plasmid pC metA_Cd (*Corynebacterium diphtheriae*) ist als SEQ ID NO:69 aufgeführt. Die entsprechende Plasmidkarte ist in Figur 3 dargestellt.

30

LOCUS pC_metA_Cd 6472 bp DNA circular
FEATURES Location/Qualifiers
CDS 313..1416
/vntifkey="4"

35

```

                    /label=metA\Corynebacterium\diphtheriae
CDS      1838..2629
                    /vntifkey="4"
                    /label=Kan\R
5  CDS      4910..6031
                    /vntifkey="4"
                    /label=Rep\Protein
CDS      3902..4576
                    /vntifkey="4"
10 CDS      complement(2896..3756)
                    /vntifkey="4"
                    /label=Ori\-EC\pMB)

```

15 **Beispiel 8:** Transformation des Stammes LU1479 lysC 311ile ET-16 mit dem Plasmid pC metA_Cd

Der Stamm LU1479 lysC 311ile ET-16 wurde mit dem Plasmid pC metA_Cd nach der beschriebenen Methode (Liebl, et al. (1989) FEMS Microbiology Letters 53:299-303) transformiert. Die Transformationsmischung wurde auf CM-Platten plattiert, die zusätzlich 20mg/l Kanamycin enthielten, um eine Selektion auf Plasmid-haltige Zellen zu erreichen. Erhaltene Kan-resistente Klone wurden gepickt und vereinzelt. Die Methionin-Produktivität der Klone wurde in einem Schüttelkolbenversuch (s. Beispiel 6) untersucht. Der Stamm LU1479 lysC 311ile ET-16 pC metA_Cd produzierte im Vergleich zu LU1479 lysC 311ile ET-16 signifikant mehr Methionin.

Patentansprüche

1. Verfahren zur fermentativen Herstellung wenigstens einer schwefelhaltigen
 5 Feinchemikalie, welches folgende Schritte umfasst:
- 10 a) Fermentation einer die gewünschte schwefelhaltige Feinchemikalie produzierenden coryneformen Bakterienkultur, wobei in den coryneformen Bakterien zumindest eine heterologe Nukleotidsequenz exprimiert wird, welche für ein Protein mit Homoserin-O-Acetyl-Transferase (metA)-Aktivität kodiert;
 - b) Anreicherung der schwefelhaltigen Feinchemikalie im Medium oder in den Zellen der Bakterien, und
 - c) Isolieren der schwefelhaltigen Feinchemikalie.
- 15 2. Verfahren nach Anspruch 1, wobei die schwefelhaltige Feinchemikalie L-Methionin umfasst.
3. Verfahren nach einem der vorhergehenden Ansprüche, wobei die heterologe metA-kodierende Nukleotidsequenz zur metA-kodierenden Sequenz aus *Corynebacterium glutamicum* ATCC 13032 eine Sequenzhomologie vom weniger als 100% aufweist.
- 20 4. Verfahren nach Anspruch 3, wobei die metA-kodierende Sequenz aus einem der folgenden Organismen abgeleitet ist:

<i>Corynebacterium diptheriae</i>	ATCC 14779
<i>Mycobacterium leprae</i>	ATCC 43910
<i>Mycobacterium tuberculosis</i> CDC1551	ATCC 25584
<i>Chlorobium tepidum</i>	ATCC 49652
<i>Pseudomonas aeruginosa</i>	ATCC 17933
<i>Caulobacter crescentus</i>	ATCC 19089
<i>Neisseria gonorrhoeae</i>	ATCC 53420
<i>Neisseria meningitidis</i>	ATCC 53414
<i>Pseudomonas fluorescens</i>	ATCC 13525
<i>Burkholderia cepacia</i>	ATCC 25416
<i>Nitrosomonas europaea</i>	ATCC 19718
<i>Haemophilus influenzae</i>	ATCC 51907
<i>Halobacterium</i> sp NRC1	ATCC 33170
<i>Thermus thermophilus</i>	ATCC 27634
<i>Deinococcus radiodurans</i>	ATCC 13939
<i>Saccharomyces cerevisiae</i>	ATCC 10751
<i>Schizosaccharomyces pombe</i>	ATCC 24969
<i>Xylella fastidiosa</i>	ATCC 35881
<i>Emericella nidulans</i>	ATCC 36104
<i>Mesorhizobium loti</i>	ATCC 35173
<i>Acremonium crysogenum</i>	ATCC 11550

Pseudomonas putida	ATCC 47054
Staphylococcus aureus	ATCC 35556

5. Verfahren nach einem der vorhergehenden Ansprüche, wobei die metA-kodierende Sequenz eine kodierende Sequenz gemäß SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, und 45 oder eine dazu homologe Nukleotidsequenz, welche für ein Protein mit metA-Aktivität kodiert, umfasst.
6. Verfahren nach einem der vorhergehenden Ansprüche, wobei die metA-kodierende Sequenz für ein Protein mit metA-Aktivität kodiert, wobei das Protein eine Aminosäuresequenz gemäß SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, und 46 oder eine dazu homologe Aminosäuresequenz, welche für ein Protein mit metA-Aktivität steht, umfasst.
7. Verfahren nach einem der vorhergehenden Ansprüche, wobei die kodierende metA-Sequenz eine in coryneformen Bakterien replizierbare oder eine stabil in das Chromosom integrierte DNA oder eine RNA ist.
8. Verfahren gemäß Anspruch 7, wobei man
 - a) einen mit einem Plasmidvektor transformierten Bakterienstamm einsetzt der wenigstens eine Kopie der kodierenden metA-Sequenz unter der Kontrolle regulativer Sequenzen trägt, oder
 - b) einen Stamm einsetzt, in dem die kodierende metA-Sequenz in das Chromosom des Bakteriums integriert wurde
9. Verfahren nach einem der vorhergehenden Ansprüche, wobei die kodierende metA-Sequenz überexprimiert wird.
10. Verfahren gemäß einem der vorhergehenden Ansprüche, wobei man Bakterien fermentiert, in denen zusätzlich wenigstens ein weiteres Gen des Biosyntheseweges der gewünschten schwefelhaltigen Feinchemikalie verstärkt ist oder derart mutiert ist, dass es durch Stoffwechselmetabolite nicht in seiner Aktivität beeinflusst wird.

11. Verfahren gemäß einem der vorhergehenden Ansprüche, wobei man Bakterien fermentiert, in denen wenigstens ein Stoffwechselweg zumindest teilweise ausgeschaltet ist, der die Bildung der gewünschten schwefelhaltigen Feinchemikalie verringert.

5 12. Verfahren gemäß einem der vorhergehenden Ansprüche, wobei man coryneforme Bakterien fermentiert, in denen gleichzeitig wenigstens eines der Gene, ausgewählt unter

- 10 a) dem für eine Aspartatkinase kodierenden Gen *lysC*,
b) dem für die Glycerinaldehyd-3-Phosphat Dehydrogenase kodierenden Gen *gap*,
c) dem für die 3-Phosphoglycerat Kinase kodierenden Gen *pgk*,
d) dem für die Pyruvat Carboxylase kodierenden Gen *pyc*,
e) dem für die Triosephosphat Isomerase kodierenden Gen *tpi*,
f) dem für die Methylentetrahydrofolat Reduktase kodierenden Gen *metF*,
g) dem für die Cystathionin-gamma-Synthase kodierenden Gen *metB*,
15 h) dem für die Cystathionin-gamma-Lyase kodierenden Gen *metC*,
i) dem für die Serin-Hydroxymethyltransferase kodierenden Gen *glyA*,
j) dem für die O-Acetylhomoserin-Sulphydrylase kodierenden Gen *metY*,
k) dem für die Vitamin B12 abhängige Methionin-Synthase kodierenden Gen *methH*,
l) dem für die Phosphoserin-Aminotransferase kodierenden Gen *serC*,
20 m) dem *serB* Gen, das für die Phosphoserin-Phosphatase kodiert,
n) dem *cysE* Gen, das für die Serine Acetyl-Transferase kodiert, und
o) dem *hom* Gen, das eine Homoserin-Dehydrogenase kodiert,

25 überexprimiert oder so mutiert ist, dass die korrespondierenden Proteine, verglichen mit nicht mutierten Proteinen, in geringerem Maße oder nicht durch Stoffwechselmetabolite in ihrer Aktivität beeinflusst werden.

13. Verfahren gemäß einem der vorhergehenden Ansprüche, wobei man coryneformen Bakterien fermentiert, in denen gleichzeitig wenigstens eines der Gene, ausgewählt unter

- 30 a) dem für die Homoserine-Kinase kodierenden Gen *thrB*,
b) dem für die Threonin Dehydratase kodierenden Gen *ilvA*,
c) dem für die Threonin Synthase kodierenden Gen *thrC*
d) dem für die Meso-Diaminopimelat D-Dehydrogenase kodierenden Gen *ddh*
e) dem für die Phosphoenolpyruvat-Carboxykinase kodierenden Gen *pck*,
35 f) dem für die Glucose-6-Phosphat-6-Isomerase kodierenden Gen *pgi*,
g) dem für die Pyruvat-Oxidase kodierenden Gen *poxB*,
h) dem für die Dihydrodipicolinat Synthase kodierenden Gen *dapA*,

- i) dem für die Dihydrodipicolinat Reduktase kodierenden Gen *dapB*; oder
- j) dem für die Diaminopicolinat Decarboxylase kodierenden Gen

durch Veränderung der Expressionsrate oder durch Einführung einer gezielten Mutation
abschwächt ist.

5

14. Verfahren gemäß einem oder mehreren der vorhergehenden Ansprüche, wobei man Mikroorganismen der Art *Corynebacterium glutamicum* einsetzt.

10

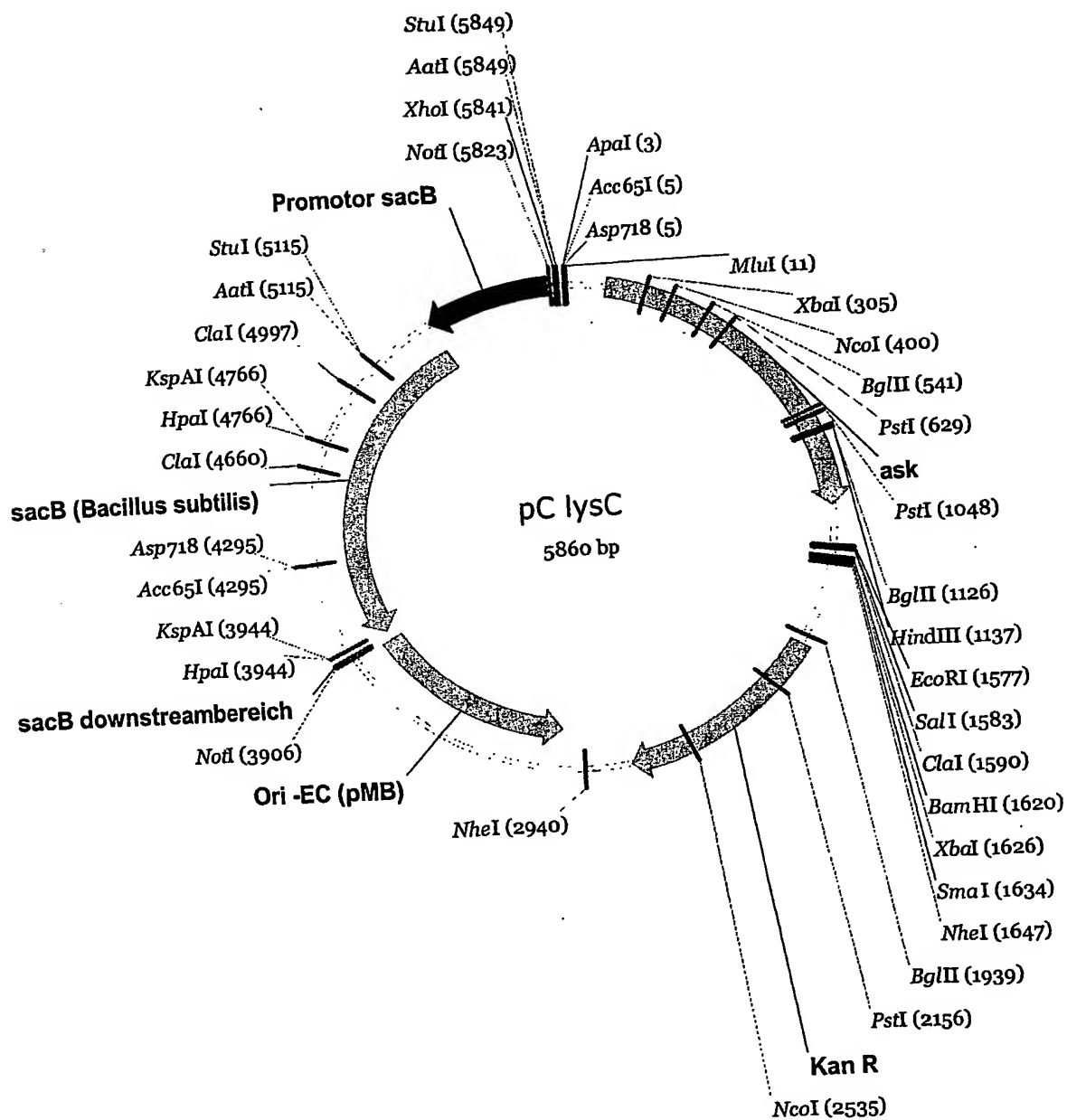
15. Verfahren zur Herstellung eines L-Methionin haltigen Tierfuttermittel-Additivs aus Fermentationsbrühen, welches folgende Schritte umfasst

- a) Kultivierung und Fermentation eines L-Methionin produzierenden Mikroorganismus in einem Fermentationsmedium;
- b) Entfernung von Wasser aus der L-Methionin haltigen Fermentationsbrühe;
- c) Entfernung der während der Fermentation gebildeten Biomasse in einer Menge von 0 bis 100 Gew.-%; und
- d) Trocknung der gemäß b) und/oder c) erhaltenen Fermentationsbrühe, um das Tierfuttermittel-Additiv in der gewünschten Pulver- oder Granulatform zu erhalten.

15

- 20 16. Verfahren gemäß Anspruch 15, wobei man Mikroorganismen gemäß der Definition in einem der Ansprüche 1 bis 14 einsetzt.

Fig. 1



2/3

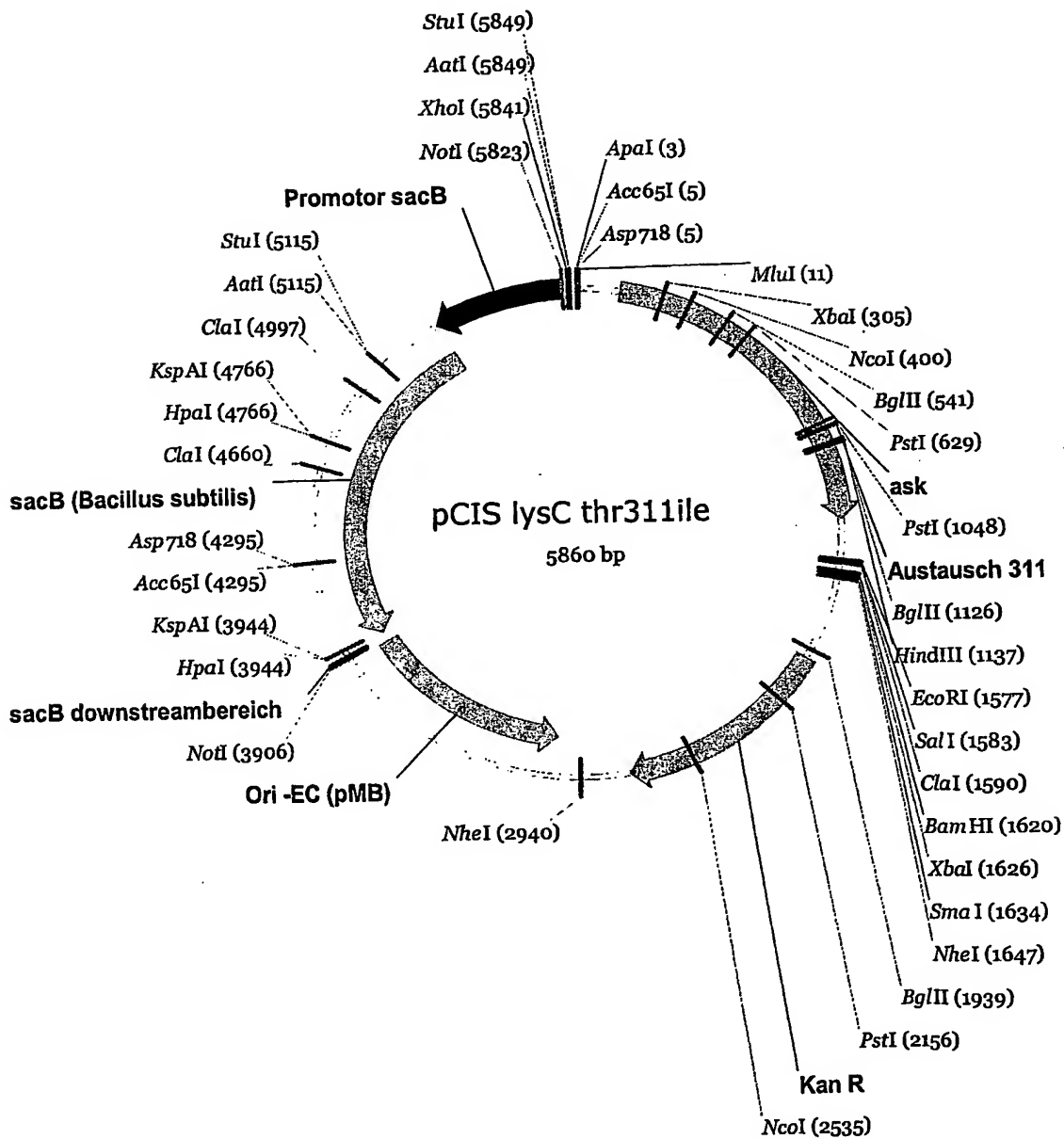
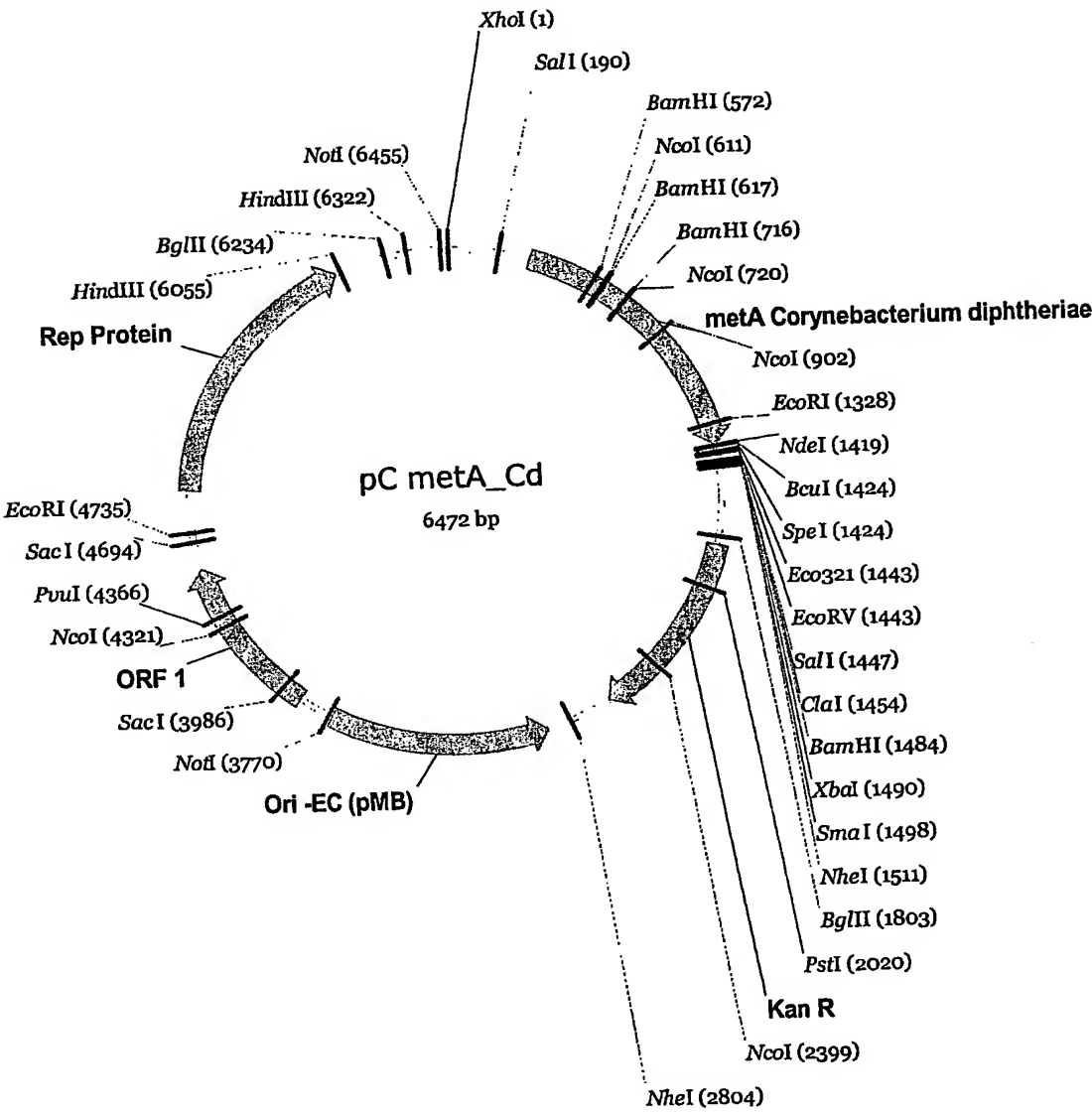
Fig. 2

Fig. 3



SEQUENZPROTOKOLL

<110> BASF Aktiengesellschaft

<120> MetA

<130> M/43127

<140>

<141>

<160> 58

<210> 1

<211> 1104

<212> DNA

<213> Corynebacterium diphtheriae

<220>

<221> CDS

<222> (1) .. (1101)

<223> RDI00386

<400> 1

atg ctc acc acc aca ggg acg ctc acg cac caa aaa atc gga gac ttt	48
Met Leu Thr Thr Thr Gly Thr Leu Thr His Gln Lys Ile Gly Asp Phe	
1 5 10 15	
tac acc gaa gcc gga gcg acg ctt cac gac gta acc atc gcc tac caa	96
Tyr Thr Glu Ala Gly Ala Thr Leu His Asp Val Thr Ile Ala Tyr Gln	
20 25 30	
gca tgg ggc cac tac acc ggc acc aat ctc atc gtt ctc gaa cat gcc	144
Ala Trp Gly His Tyr Thr Gly Thr Asn Leu Ile Val Leu Glu His Ala	
35 40 45	
ctg acc ggc gac tct aac gct att tca tgg tgg gac gga ctg att ggc	192
Leu Thr Gly Asp Ser Asn Ala Ile Ser Trp Trp Asp Gly Leu Ile Gly	
50 55 60	
cct ggc aaa gca ctc gac acc aac cgc tac tgc atc cta tgc acc aac	240
Pro Gly Lys Ala Leu Asp Thr Asn Arg Tyr Cys Ile Leu Cys Thr Asn	
65 70 75 80	
gtg ctc gga gga tgc aaa gga tcc acc gga ccg agc agt cca cac cca	288
Val Leu Gly Gly Cys Lys Gly Ser Thr Gly Pro Ser Ser Pro His Pro	
85 90 95	
gac gga aaa cca tgg gga tcc aga ttt cca gcc ctt tca atc cgt gac	336
Asp Gly Lys Pro Trp Gly Ser Arg Phe Pro Ala Leu Ser Ile Arg Asp	
100 105 110	
ctt gtc aat gcc gaa aaa caa ctt ttc gac cac ctc ggc atc aat aaa	384
Leu Val Asn Ala Glu Lys Gln Leu Phe Asp His Leu Gly Ile Asn Lys	
115 120 125	
att cac gca atc atc ggc gga tcc atg gga ggc gca cgc acc ctc gaa	432
Ile His Ala Ile Ile Gly Gly Ser Met Gly Gly Ala Arg Thr Leu Glu	
130 135 140	
tgg gct gca ctc cac cca cac atg atg acg act gga ttc gtc ata gca	480
Trp Ala Ala Leu His Pro His Met Met Thr Thr Gly Phe Val Ile Ala	

145	150	155	160	
gtc tca gca cgc gca agc gct tgg caa atc ggt att caa act gca caa				528
Val Ser Ala Arg Ala Ser Ala Trp Gln Ile Gly Ile Gln Thr Ala Gln				
	165	170	175	
atc agc gcc ata gaa ctc gac ccc cac tgg aac ggc ggc gat tac tac				576
Ile Ser Ala Ile Glu Leu Asp Pro His Trp Asn Gly Gly Asp Tyr Tyr				
	180	185	190	
agc ggt cac gca cca tgg gaa gga atc gcc gcc gct cgc cgg atc gcc				624
Ser Gly His Ala Pro Trp Glu Gly Ile Ala Ala Ala Arg Arg Ile Ala				
	195	200	205	
cac ctc acc tat cgc ggc gaa cta gaa ata gac gaa cga ttc ggc act				672
His Leu Thr Tyr Arg Gly Glu Leu Glu Ile Asp Glu Arg Phe Gly Thr				
	210	215	220	
tcc gca caa cac ggt gaa aac cca ctc ggc ccc ttc cga gat cca cat				720
Ser Ala Gln His Gly Glu Asn Pro Leu Gly Pro Phe Arg Asp Pro His				
	225	230	235	240
caa cgt ttt gcg gtc acg agc tac ctc caa cac caa ggc atc aaa ctc				768
Gln Arg Phe Ala Val Thr Ser Tyr Leu Gln His Gln Gly Ile Lys Leu				
	245	250	255	
gct caa cga ttc gat gca ggt agt tac gtc gtg ctt acc gaa gcc ctc				816
Ala Gln Arg Phe Asp Ala Gly Ser Tyr Val Val Leu Thr Glu Ala Leu				
	260	265	270	
aat cgt cat gac atc gga cgc ggc cga ggc gga ctc aac aaa gcc ctc				864
Asn Arg His Asp Ile Gly Arg Gly Arg Gly Gly Leu Asn Lys Ala Leu				
	275	280	285	
agc gca atc aca gtc ccc atc atg att gct ggc gtt gat acc gat att				912
Ser Ala Ile Thr Val Pro Ile Met Ile Ala Gly Val Asp Thr Asp Ile				
	290	295	300	
ctc tac ccc tat cac cag caa gaa cac cta tca cga aat cta ggc aac				960
Leu Tyr Pro Tyr His Gln Gln Glu His Leu Ser Arg Asn Leu Gly Asn				
	305	310	315	320
cta ctc gct atg gca aaa atc agc tca cca gta ggc cac gac gct ttc				1008
Leu Leu Ala Met Ala Lys Ile Ser Ser Pro Val Gly His Asp Ala Phe				
	325	330	335	
ctc aca gaa ttc cga caa atg gag cga atc cta aga cat ttc atg gag				1056
Leu Thr Glu Phe Arg Gln Met Glu Arg Ile Leu Arg His Phe Met Glu				
	340	345	350	
ctt tcg gaa gga atc gac gat tcc ttc cga acc aaa cta gag cgc				1101
Leu Ser Glu Gly Ile Asp Asp Ser Phe Arg Thr Lys Leu Glu Arg				
	355	360	365	
tga				1104

<210> 2

<211> 367

<212> PRT

<213> Corynebacterium diphtheriae

<400> 2

Met Leu Thr Thr Thr Gly Thr Leu Thr His Gln Lys Ile Gly Asp Phe
 1 5 10 15
 Tyr Thr Glu Ala Gly Ala Thr Leu His Asp Val Thr Ile Ala Tyr Gln
 20 25 30
 Ala Trp Gly His Tyr Thr Gly Thr Asn Leu Ile Val Leu Glu His Ala
 35 40 45
 Leu Thr Gly Asp Ser Asn Ala Ile Ser Trp Trp Asp Gly Leu Ile Gly
 50 55 60
 Pro Gly Lys Ala Leu Asp Thr Asn Arg Tyr Cys Ile Leu Cys Thr Asn
 65 70 75 80
 Val Leu Gly Gly Cys Lys Gly Ser Thr Gly Pro Ser Ser Pro His Pro
 85 90 95
 Asp Gly Lys Pro Trp Gly Ser Arg Phe Pro Ala Leu Ser Ile Arg Asp
 100 105 110
 Leu Val Asn Ala Glu Lys Gln Leu Phe Asp His Leu Gly Ile Asn Lys
 115 120 125
 Ile His Ala Ile Ile Gly Gly Ser Met Gly Gly Ala Arg Thr Leu Glu
 130 135 140
 Trp Ala Ala Leu His Pro His Met Met Thr Thr Gly Phe Val Ile Ala
 145 150 155 160
 Val Ser Ala Arg Ala Ser Ala Trp Gln Ile Gly Ile Gln Thr Ala Gln
 165 170 175
 Ile Ser Ala Ile Glu Leu Asp Pro His Trp Asn Gly Gly Asp Tyr Tyr
 180 185 190
 Ser Gly His Ala Pro Trp Glu Gly Ile Ala Ala Ala Arg Arg Ile Ala
 195 200 205
 His Leu Thr Tyr Arg Gly Glu Leu Glu Ile Asp Glu Arg Phe Gly Thr
 210 215 220
 Ser Ala Gln His Gly Glu Asn Pro Leu Gly Pro Phe Arg Asp Pro His
 225 230 235 240
 Gln Arg Phe Ala Val Thr Ser Tyr Leu Gln His Gln Gly Ile Lys Leu
 245 250 255
 Ala Gln Arg Phe Asp Ala Gly Ser Tyr Val Val Leu Thr Glu Ala Leu
 260 265 270
 Asn Arg His Asp Ile Gly Arg Gly Arg Gly Gly Leu Asn Lys Ala Leu
 275 280 285
 Ser Ala Ile Thr Val Pro Ile Met Ile Ala Gly Val Asp Thr Asp Ile
 290 295 300
 Leu Tyr Pro Tyr His Gln Gln Glu His Leu Ser Arg Asn Leu Gly Asn
 305 310 315 320
 Leu Leu Ala Met Ala Lys Ile Ser Ser Pro Val Gly His Asp Ala Phe
 325 330 335

Leu Thr Glu Phe Arg Gln Met Glu Arg Ile Leu Arg His Phe Met Glu
 340 345 350

Leu Ser Glu Gly Ile Asp Asp Ser Phe Arg Thr Lys Leu Glu Arg
 355 360 365

<210> 3

<211> 1149

<212> DNA

<213> Mycobacterium leprae

<220>

<221> CDS

<222> (1) .. (1146)

<223> RML02951

<220>

<221> unsure

<222> 224 .. 224

<223> All occurrences of n indicate any nucleotide

<400> 3

atg aca atc tcc aag gtc cct acc cag aag ctg ccg gcc gaa ggc gag 48
 Met Thr Ile Ser Lys Val Pro Thr Gln Lys Leu Pro Ala Glu Gly Glu
 1 5 10 15

gtc ggc ttg gtc gac atc ggc tca ctt acc acc gaa agc ggt gcc gtc 96
 Val Gly Leu Val Asp Ile Gly Ser Leu Thr Thr Glu Ser Gly Ala Val
 20 25 30

atc gac gat gtc tgc atc gcc gtt cag cgc tgg ggg gaa ttg tcg ccc 144
 Ile Asp Asp Val Cys Ile Ala Val Gln Arg Trp Gly Glu Leu Ser Pro
 35 40 45

acg cga gac aac gta gtg atg gta ctg cat gca ctc acc ggt gac tcg 192
 Thr Arg Asp Asn Val Val Met Val Leu His Ala Leu Thr Gly Asp Ser
 50 55 60

cac atc acc ggg ccc gcc gga ccg gga cat cnc aca ccc ggc tgg tgg 240
 His Ile Thr Gly Pro Ala Gly Pro Gly His Xaa Thr Pro Gly Trp Trp
 65 70 75 80

gac tgg ata gct gga ccg ggt gca cca atc gac acc aac cgc tgg tgc 288
 Asp Trp Ile Ala Gly Pro Gly Ala Pro Ile Asp Thr Asn Arg Trp Cys
 85 90 95

gcg ata gcc acc aac gtg ctg ggc ggt tgc cgt ggc tcc acc ggc cct 336
 Ala Ile Ala Thr Asn Val Leu Gly Gly Cys Arg Gly Ser Thr Gly Pro
 100 105 110

agt tcg ctt gcc cgc gac gga aag cct tgg ggt tca aga ttt ccg ctg 384
 Ser Ser Leu Ala Arg Asp Gly Lys Pro Trp Gly Ser Arg Phe Pro Leu
 115 120 125

ata tct ata cgc gac cag gta gag gca gat atc gct gca ctg gcc gcc 432
 Ile Ser Ile Arg Asp Gln Val Glu Ala Asp Ile Ala Ala Leu Ala Ala
 130 135 140

atg gga att aca aag gtt gcc gcc gtc gtt gga gga tct atg ggc ggg 480
 Met Gly Ile Thr Lys Val Ala Ala Val Val Gly Gly Ser Met Gly Gly

145	150	155	160	
gcg cgt gca ctg gaa tgg atc atc ggc cac ccg gac caa gtc cgg gcc				528
Ala Arg Ala Leu Glu Trp Ile Ile Gly His Pro Asp Gln Val Arg Ala				
	165	170	175	
ggg ctg ttg ctg gcg gtc ggt gtg cgc gcc acc gcc gac cag atc ggc				576
Gly Leu Leu Leu Ala Val Gly Val Arg Ala Thr Ala Asp Gln Ile Gly				
	180	185	190	
acc caa acc acc caa atc gca gcc atc aag aca gac ccg aac tgg caa				624
Thr Gln Thr Thr Gln Ile Ala Ala Ile Lys Thr Asp Pro Asn Trp Gln				
	195	200	205	
ggc ggt gac tac tac gag aca ggg agg gca cca gag aac ggc ttg aca				672
Gly Gly Asp Tyr Tyr Glu Thr Gly Arg Ala Pro Glu Asn Gly Leu Thr				
	210	215	220	
att gcc cgc cgc ttc gcc cac ctg acc tac cgc agc gag gtc gag ctc				720
Ile Ala Arg Arg Phe Ala His Leu Thr Tyr Arg Ser Glu Val Glu Leu				
	225	230	235	240
gac acc cgg ttt gcc aac aac aac caa ggc aat gag gac ccg gcg acg				768
Asp Thr Arg Phe Ala Asn Asn Asn Gln Gly Asn Glu Asp Pro Ala Thr				
	245	250	255	
ggc ggg cgt tac gca gtg cag agt tac cta gag cac cag ggt gac aag				816
Gly Gly Arg Tyr Ala Val Gln Ser Tyr Leu Glu His Gln Gly Asp Lys				
	260	265	270	
cta ttg gcc cgc ttt gac gca ggc agc tac gtg gtc ttg acc gaa acg				864
Leu Leu Ala Arg Phe Asp Ala Gly Ser Tyr Val Val Leu Thr Glu Thr				
	275	280	285	
ctg aac agc cac gac gtt ggc cgg ggc cgc gga ggg atc ggt aca gcg				912
Leu Asn Ser His Asp Val Gly Arg Gly Arg Gly Gly Ile Gly Thr Ala				
	290	295	300	
ctg cgc ggg tgc ccg gta ccg gtg gtg gtg ggt ggc att acc tcg gat				960
Leu Arg Gly Cys Pro Val Pro Val Val Val Gly Gly Ile Thr Ser Asp				
	305	310	315	320
cgg ctc tac cca ctg cgc ttg cag cag gag ctg gcc gag atg ctg ccg				1008
Arg Leu Tyr Pro Leu Arg Leu Gln Gln Glu Leu Ala Glu Met Leu Pro				
	325	330	335	
ggc tgc acc ggg ctg cag gtt gta gac tcc acc tac ggg cac gac ggc				1056
Gly Cys Thr Gly Leu Gln Val Val Asp Ser Thr Tyr Gly His Asp Gly				
	340	345	350	
ttc ctg gtg gaa tcc gag gcc gtc ggc aaa ttg atc cgt caa acc ctc				1104
Phe Leu Val Glu Ser Glu Ala Val Gly Lys Leu Ile Arg Gln Thr Leu				
	355	360	365	
gaa ttg gcc gac gtg ggt tcc aag gaa gac gcg tgt tcg caa				1146
Glu Leu Ala Asp Val Gly Ser Lys Glu Asp Ala Cys Ser Gln				
	370	375	380	
tga				1149

<211> 382

<212> PRT

<213> Mycobacterium leprae

<220>

<221> unsure

<222> 75 .. 75

<223> All occurrences of Xaa indicate any amino acid

<400> 4

```

Met Thr Ile Ser Lys Val Pro Thr Gln Lys Leu Pro Ala Glu Gly Glu
 1             5             10             15

Val Gly Leu Val Asp Ile Gly Ser Leu Thr Thr Glu Ser Gly Ala Val
      20             25             30

Ile Asp Asp Val Cys Ile Ala Val Gln Arg Trp Gly Glu Leu Ser Pro
      35             40             45

Thr Arg Asp Asn Val Val Met Val Leu His Ala Leu Thr Gly Asp Ser
      50             55             60

His Ile Thr Gly Pro Ala Gly Pro Gly His Xaa Thr Pro Gly Trp Trp
      65             70             75             80

Asp Trp Ile Ala Gly Pro Gly Ala Pro Ile Asp Thr Asn Arg Trp Cys
      85             90             95

Ala Ile Ala Thr Asn Val Leu Gly Gly Cys Arg Gly Ser Thr Gly Pro
      100             105             110

Ser Ser Leu Ala Arg Asp Gly Lys Pro Trp Gly Ser Arg Phe Pro Leu
      115             120             125

Ile Ser Ile Arg Asp Gln Val Glu Ala Asp Ile Ala Ala Leu Ala Ala
      130             135             140

Met Gly Ile Thr Lys Val Ala Ala Val Val Gly Gly Ser Met Gly Gly
      145             150             155             160

Ala Arg Ala Leu Glu Trp Ile Ile Gly His Pro Asp Gln Val Arg Ala
      165             170             175

Gly Leu Leu Leu Ala Val Gly Val Arg Ala Thr Ala Asp Gln Ile Gly
      180             185             190

Thr Gln Thr Thr Gln Ile Ala Ala Ile Lys Thr Asp Pro Asn Trp Gln
      195             200             205

Gly Gly Asp Tyr Tyr Glu Thr Gly Arg Ala Pro Glu Asn Gly Leu Thr
      210             215             220

Ile Ala Arg Arg Phe Ala His Leu Thr Tyr Arg Ser Glu Val Glu Leu
      225             230             235             240

Asp Thr Arg Phe Ala Asn Asn Asn Gln Gly Asn Glu Asp Pro Ala Thr
      245             250             255

Gly Gly Arg Tyr Ala Val Gln Ser Tyr Leu Glu His Gln Gly Asp Lys
      260             265             270

Leu Leu Ala Arg Phe Asp Ala Gly Ser Tyr Val Val Leu Thr Glu Thr
      275             280             285

```

Leu Asn Ser His Asp Val Gly Arg Gly Arg Gly Gly Ile Gly Thr Ala
 290 295 300
 Leu Arg Gly Cys Pro Val Pro Val Val Val Gly Gly Ile Thr Ser Asp
 305 310 315 320
 Arg Leu Tyr Pro Leu Arg Leu Gln Gln Glu Leu Ala Glu Met Leu Pro
 325 330 335
 Gly Cys Thr Gly Leu Gln Val Val Asp Ser Thr Tyr Gly His Asp Gly
 340 345 350
 Phe Leu Val Glu Ser Glu Ala Val Gly Lys Leu Ile Arg Gln Thr Leu
 355 360 365
 Glu Leu Ala Asp Val Gly Ser Lys Glu Asp Ala Cys Ser Gln
 370 375 380

<210> 5

<211> 1140

<212> DNA

<213> Mycobacterium tuberculosis

<220>

<221> CDS

<222> (1)..(1137)

<223> RMTB03565

<400> 5

atg acg atc tcc gat gta ccc acc cag acg ctg ccc gcc gaa ggc gaa 48
 Met Thr Ile Ser Asp Val Pro Thr Gln Thr Leu Pro Ala Glu Gly Glu
 1 5 10 15
 atc ggc ctg ata gac gtc ggc tcg ctg caa ctg gaa agc ggg gcg gtg 96
 Ile Gly Leu Ile Asp Val Gly Ser Leu Gln Leu Glu Ser Gly Ala Val
 20 25 30
 atc gac gat gtc tgt atc gcc gtg caa cgc tgg ggc aaa ttg tcg ccc 144
 Ile Asp Asp Val Cys Ile Ala Val Gln Arg Trp Gly Lys Leu Ser Pro
 35 40 45
 gca cgg gac aac gtg gtg gtg gtc ttg cac gcg ctc acc ggc gac tcg 192
 Ala Arg Asp Asn Val Val Val Leu His Ala Leu Thr Gly Asp Ser
 50 55 60
 cac atc act gga ccc gcc gga ccc ggc cac ccc acc ccc ggc tgg tgg 240
 His Ile Thr Gly Pro Ala Gly Pro Gly His Pro Thr Pro Gly Trp Trp
 65 70 75 80
 gac ggg gtg gcc ggg ccg agt gcg ccg att gac acc acc cgc tgg tgc 288
 Asp Gly Val Ala Gly Pro Ser Ala Pro Ile Asp Thr Thr Arg Trp Cys
 85 90 95
 gcg gta gct acc aat gtg ctc ggc ggc tgc cgc ggc tcc acc ggg ccc 336
 Ala Val Ala Thr Asn Val Leu Gly Gly Cys Arg Gly Ser Thr Gly Pro
 100 105 110
 agc tcg ctt gcc cgc gac gga aag cct tgg ggc tca aga ttt ccg ctg 384
 Ser Ser Leu Ala Arg Asp Gly Lys Pro Trp Gly Ser Arg Phe Pro Leu
 115 120 125

atc tgc ata cgt gac cag gtg cag gcg gac gtc gcg gcg ctg gcc gcg	432
Ile Ser Ile Arg Asp Gln Val Gln Ala Asp Val Ala Ala Leu Ala Ala	
130 135 140	
ctg ggc atc acc gag gtc gcc gcc gtc gtc ggc ggc tcc atg ggc ggc	480
Leu Gly Ile Thr Glu Val Ala Ala Val Val Gly Gly Ser Met Gly Gly	
145 150 155 160	
gcc cgg gcc ctg gaa tgg gtg gtc ggc tac ccg gat cgg gtc cga gcc	528
Ala Arg Ala Leu Glu Trp Val Val Gly Tyr Pro Asp Arg Val Arg Ala	
165 170 175	
gga ttg ctg ctg gcg gtc ggt gcg cgt gcc acc gca gac cag atc ggc	576
Gly Leu Leu Leu Ala Val Gly Ala Arg Ala Thr Ala Asp Gln Ile Gly	
180 185 190	
acg cag aca acg caa atc gcg gcc atc aaa gcc gac ccg gac tgg cag	624
Thr Gln Thr Thr Gln Ile Ala Ala Ile Lys Ala Asp Pro Asp Trp Gln	
195 200 205	
agc ggc gac tac cac gag acg ggg agg gca cca gac gcc ggg ctg cga	672
Ser Gly Asp Tyr His Glu Thr Gly Arg Ala Pro Asp Ala Gly Leu Arg	
210 215 220	
ctc gcc cgc cgc ttc gcg cac ctc acc tac cgc ggc gag atc gag ctc	720
Leu Ala Arg Arg Phe Ala His Leu Thr Tyr Arg Gly Glu Ile Glu Leu	
225 230 235 240	
gac acc cgg ttc gcc aac cac aac cag ggc aac gag gat ccg acg gcc	768
Asp Thr Arg Phe Ala Asn His Asn Gln Gly Asn Glu Asp Pro Thr Ala	
245 250 255	
ggc ggg cgc tac gcg gtg caa agt tat ctg gaa cac caa gga gac aaa	816
Gly Gly Arg Tyr Ala Val Gln Ser Tyr Leu Glu His Gln Gly Asp Lys	
260 265 270	
ctg tta tcc cgg ttc gac gcc ggc agc tac gtg att ctc acc gag gcg	864
Leu Leu Ser Arg Phe Asp Ala Gly Ser Tyr Val Ile Leu Thr Glu Ala	
275 280 285	
ctc aac agc cac gac gtc ggc cgc ggc cgc ggc ggc gtc tcc gcg gct	912
Leu Asn Ser His Asp Val Gly Arg Gly Arg Gly Gly Val Ser Ala Ala	
290 295 300	
ctg cgc gcc tgc ccg gtg ccg gtg gtg gtg ggc ggc atc acc tcc gac	960
Leu Arg Ala Cys Pro Val Pro Val Val Val Gly Gly Ile Thr Ser Asp	
305 310 315 320	
cgg ctc tac ccg ctg cgc ctg cag cag gag ctg gcc gac ctg ctg ccg	1008
Arg Leu Tyr Pro Leu Arg Leu Gln Gln Glu Leu Ala Asp Leu Leu Pro	
325 330 335	
ggc tgc gcc ggg ctg cga gtc gtc gag tgc gtc tac gga cac gac ggc	1056
Gly Cys Ala Gly Leu Arg Val Val Glu Ser Val Tyr Gly His Asp Gly	
340 345 350	
ttc ctg gtg gaa acc gag gcc gtg ggc gaa ttg atc cgc cag aca ctg	1104
Phe Leu Val Glu Thr Glu Ala Val Gly Glu Leu Ile Arg Gln Thr Leu	
355 360 365	

gga ttg gct gat cgt gaa ggc gcg tgt cgg cgg tga
 Gly Leu Ala Asp Arg Glu Gly Ala Cys Arg Arg
 370 375

1140

<210> 6

<211> 379

<212> PRT

<213> Mycobacterium tuberculosis

<400> 6

Met Thr Ile Ser Asp Val Pro Thr Gln Thr Leu Pro Ala Glu Gly Glu
 1 5 10 15

Ile Gly Leu Ile Asp Val Gly Ser Leu Gln Leu Glu Ser Gly Ala Val
 20 25 30

Ile Asp Asp Val Cys Ile Ala Val Gln Arg Trp Gly Lys Leu Ser Pro
 35 40 45

Ala Arg Asp Asn Val Val Val Val Leu His Ala Leu Thr Gly Asp Ser
 50 55 60

His Ile Thr Gly Pro Ala Gly Pro Gly His Pro Thr Pro Gly Trp Trp
 65 70 75 80

Asp Gly Val Ala Gly Pro Ser Ala Pro Ile Asp Thr Thr Arg Trp Cys
 85 90 95

Ala Val Ala Thr Asn Val Leu Gly Gly Cys Arg Gly Ser Thr Gly Pro
 100 105 110

Ser Ser Leu Ala Arg Asp Gly Lys Pro Trp Gly Ser Arg Phe Pro Leu
 115 120 125

Ile Ser Ile Arg Asp Gln Val Gln Ala Asp Val Ala Ala Leu Ala Ala
 130 135 140

Leu Gly Ile Thr Glu Val Ala Ala Val Val Gly Gly Ser Met Gly Gly
 145 150 155 160

Ala Arg Ala Leu Glu Trp Val Val Gly Tyr Pro Asp Arg Val Arg Ala
 165 170 175

Gly Leu Leu Leu Ala Val Gly Ala Arg Ala Thr Ala Asp Gln Ile Gly
 180 185 190

Thr Gln Thr Thr Gln Ile Ala Ala Ile Lys Ala Asp Pro Asp Trp Gln
 195 200 205

Ser Gly Asp Tyr His Glu Thr Gly Arg Ala Pro Asp Ala Gly Leu Arg
 210 215 220

Leu Ala Arg Arg Phe Ala His Leu Thr Tyr Arg Gly Glu Ile Glu Leu
 225 230 235 240

Asp Thr Arg Phe Ala Asn His Asn Gln Gly Asn Glu Asp Pro Thr Ala
 245 250 255

Gly Gly Arg Tyr Ala Val Gln Ser Tyr Leu Glu His Gln Gly Asp Lys
 260 265 270

Leu Leu Ser Arg Phe Asp Ala Gly Ser Tyr Val Ile Leu Thr Glu Ala

275 280 285
 Leu Asn Ser His Asp Val Gly Arg Gly Arg Gly Gly Val Ser Ala Ala
 290 295 300
 Leu Arg Ala Cys Pro Val Pro Val Val Val Gly Gly Ile Thr Ser Asp
 305 310 315 320
 Arg Leu Tyr Pro Leu Arg Leu Gln Gln Glu Leu Ala Asp Leu Leu Pro
 325 330 335
 Gly Cys Ala Gly Leu Arg Val Val Glu Ser Val Tyr Gly His Asp Gly
 340 345 350
 Phe Leu Val Glu Thr Glu Ala Val Gly Glu Leu Ile Arg Gln Thr Leu
 355 360 365
 Gly Leu Ala Asp Arg Glu Gly Ala Cys Arg Arg
 370 375

<210> 7

<211> 972

<212> DNA

<213> Chlorobium tepidum

<220>

<221> CDS

<222> (1) .. (969)

<223> RCL01447

<400> 7

gtg agg gtc gct tac cgt acc tgg ggt acg cta aac gca gag aaa agc 48
 Val Arg Val Ala Tyr Arg Thr Trp Gly Thr Leu Asn Ala Glu Lys Ser
 1 5 10 15
 aac gtg att ctg gtc tgc cac gcg ctg acc ggc aac gcc gac gcc gac 96
 Asn Val Ile Leu Val Cys His Ala Leu Thr Gly Asn Ala Asp Ala Asp
 20 25 30
 agc tgg tgg tgc ggc atg ttc ggt gag gga cgg gcg ttc gac gag act 144
 Ser Trp Trp Cys Gly Met Phe Gly Glu Gly Arg Ala Phe Asp Glu Thr
 35 40 45
 cgg gac ttc atc gta tgc agc aac gtg ctt gga agc tgc tac gga acg 192
 Arg Asp Phe Ile Val Cys Ser Asn Val Leu Gly Ser Cys Tyr Gly Thr
 50 55 60
 acc ggg ccg atg tcg gtg aat ccg ctg agt ggc agg cac tac ggt ccc 240
 Thr Gly Pro Met Ser Val Asn Pro Leu Ser Gly Arg His Tyr Gly Pro
 65 70 75 80
 gat ttt ccg cgc att acc att cgc gac atg gtg aat gtt cag cga tta 288
 Asp Phe Pro Arg Ile Thr Ile Arg Asp Met Val Asn Val Gln Arg Leu
 85 90 95
 ttg ctt cgt tcg ctc ggc atc gac cgg atc cgg ctc atc gtt ggt gca 336
 Leu Leu Arg Ser Leu Gly Ile Asp Arg Ile Arg Leu Ile Val Gly Ala
 100 105 110
 tcg ctt ggc ggg atg cag gtg ctc gaa tgg ggc gca atg tat ccc gaa 384
 Ser Leu Gly Gly Met Gln Val Leu Glu Trp Gly Ala Met Tyr Pro Glu
 115 120 125

atg gcc ggg gcg ctg atg ccg atg ggc gtt tcg ggt cgt cat tcg gcg 432
 Met Ala Gly Ala Leu Met Pro Met Gly Val Ser Gly Arg His Ser Ala
 130 135 140

tgg tgc atc gcg cag agc gag gcg cag cgg cag gct atc gcc gcc gat 480
 Trp Cys Ile Ala Gln Ser Glu Ala Gln Arg Gln Ala Ile Ala Ala Asp
 145 150 155 160

gcg gag tgg caa gat ggc tgg tat gat ccg gag gtg cag cca cgc aaa 528
 Ala Glu Trp Gln Asp Gly Trp Tyr Asp Pro Glu Val Gln Pro Arg Lys
 165 170 175

gga ctt gcc gcc gcg cgg atg atg gcg atg tgc acc tac cgc tgc ttc 576
 Gly Leu Ala Ala Ala Arg Met Met Ala Met Cys Thr Tyr Arg Cys Phe
 180 185 190

gag aac tac cag caa cgc ttt ggc cgc aag cag cgc gag gac ggc ttg 624
 Glu Asn Tyr Gln Gln Arg Phe Gly Arg Lys Gln Arg Glu Asp Gly Leu
 195 200 205

ttc gaa gcc gaa agc tac gtg cgt cac cag ggc gac aag ctg gtt ggg 672
 Phe Glu Ala Glu Ser Tyr Val Arg His Gln Gly Asp Lys Leu Val Gly
 210 215 220

cgc ttt gat gca aac acc tat atc acg ctc acc aga gcg atg gac atg 720
 Arg Phe Asp Ala Asn Thr Tyr Ile Thr Leu Thr Arg Ala Met Asp Met
 225 230 235 240

cac gac ctc ggg cgc gga cgc gac tcc tac gaa gcg gcg ctc gga gcg 768
 His Asp Leu Gly Arg Gly Arg Asp Ser Tyr Glu Ala Ala Leu Gly Ala
 245 250 255

ctg aag atg ccg gtc gag att ctc tcc atc gac tcg gac gtg ctc tat 816
 Leu Lys Met Pro Val Glu Ile Leu Ser Ile Asp Ser Asp Val Leu Tyr
 260 265 270

ccc agg cag gag cag gag gaa ctt gcc cgc ctc att ccc ggc tca cgc 864
 Pro Arg Gln Glu Gln Glu Glu Leu Ala Arg Leu Ile Pro Gly Ser Arg
 275 280 285

ctg ctt ttc ctt gac gaa ccc tat ggc cac gac gcc ttt ctt atc gac 912
 Leu Leu Phe Leu Asp Glu Pro Tyr Gly His Asp Ala Phe Leu Ile Asp
 290 295 300

acc gag acc gtc agc cgc atg gtc tgc gag ttc aag agg cag ttg ata 960
 Thr Glu Thr Val Ser Arg Met Val Cys Glu Phe Lys Arg Gln Leu Ile
 305 310 315 320

gtt gac aat tga 972
 Val Asp Asn

<210> 8

<211> 323

<212> PRT

<213> Chlorobium tepidum

<400> 8

Val Arg Val Ala Tyr Arg Thr Trp Gly Thr Leu Asn Ala Glu Lys Ser
 1 5 10 15

Asn Val Ile Leu Val Cys His Ala Leu Thr Gly Asn Ala Asp Ala Asp
 20 25 30
 Ser Trp Trp Cys Gly Met Phe Gly Glu Gly Arg Ala Phe Asp Glu Thr
 35 40 45
 Arg Asp Phe Ile Val Cys Ser Asn Val Leu Gly Ser Cys Tyr Gly Thr
 50 55 60
 Thr Gly Pro Met Ser Val Asn Pro Leu Ser Gly Arg His Tyr Gly Pro
 65 70 75 80
 Asp Phe Pro Arg Ile Thr Ile Arg Asp Met Val Asn Val Gln Arg Leu
 85 90 95
 Leu Leu Arg Ser Leu Gly Ile Asp Arg Ile Arg Leu Ile Val Gly Ala
 100 105 110
 Ser Leu Gly Gly Met Gln Val Leu Glu Trp Gly Ala Met Tyr Pro Glu
 115 120 125
 Met Ala Gly Ala Leu Met Pro Met Gly Val Ser Gly Arg His Ser Ala
 130 135 140
 Trp Cys Ile Ala Gln Ser Glu Ala Gln Arg Gln Ala Ile Ala Ala Asp
 145 150 155 160
 Ala Glu Trp Gln Asp Gly Trp Tyr Asp Pro Glu Val Gln Pro Arg Lys
 165 170 175
 Gly Leu Ala Ala Ala Arg Met Met Ala Met Cys Thr Tyr Arg Cys Phe
 180 185 190
 Glu Asn Tyr Gln Gln Arg Phe Gly Arg Lys Gln Arg Glu Asp Gly Leu
 195 200 205
 Phe Glu Ala Glu Ser Tyr Val Arg His Gln Gly Asp Lys Leu Val Gly
 210 215 220
 Arg Phe Asp Ala Asn Thr Tyr Ile Thr Leu Thr Arg Ala Met Asp Met
 225 230 235 240
 His Asp Leu Gly Arg Gly Arg Asp Ser Tyr Glu Ala Ala Leu Gly Ala
 245 250 255
 Leu Lys Met Pro Val Glu Ile Leu Ser Ile Asp Ser Asp Val Leu Tyr
 260 265 270
 Pro Arg Gln Glu Gln Glu Glu Leu Ala Arg Leu Ile Pro Gly Ser Arg
 275 280 285
 Leu Leu Phe Leu Asp Glu Pro Tyr Gly His Asp Ala Phe Leu Ile Asp
 290 295 300
 Thr Glu Thr Val Ser Arg Met Val Cys Glu Phe Lys Arg Gln Leu Ile
 305 310 315 320
 Val Asp Asn

<211> 1149

<212> DNA

<213> *Caulobacter crescentus*

<220>

<221> CDS

<222> (1) .. (1146)

<223> RCO00727

<400> 9

atg gct gcg ctc gat ccg atc acg ccc gcc ggc ggg gga acc tgg cgg	48
Met Ala Ala Leu Asp Pro Ile Thr Pro Ala Gly Gly Gly Thr Trp Arg	
1 5 10 15	
ttt cct gcg aat gaa cct ctg cgg ctg gac tcc gga ggc gtc atc gaa	96
Phe Pro Ala Asn Glu Pro Leu Arg Leu Asp Ser Gly Gly Val Ile Glu	
20 25 30	
ggg ctg gaa atc gcc tac cag acc tac ggc cag ctg aac gcg gac aag	144
Gly Leu Glu Ile Ala Tyr Gln Thr Tyr Gly Gln Leu Asn Ala Asp Lys	
35 40 45	
tcc aac gcc gtc ctg atc tgc cac gcc ctg acg ggc gac cag cat gtg	192
Ser Asn Ala Val Leu Ile Cys His Ala Leu Thr Gly Asp Gln His Val	
50 55 60	
gcc tcg ccc cac ccc acc acc ggc aag ccc ggc tgg tgg caa cgc ctt	240
Ala Ser Pro His Pro Thr Thr Gly Lys Pro Gly Trp Trp Gln Arg Leu	
65 70 75 80	
gtt ggt ccc ggt aag ccg ctg gat ccc gcg cgg cac ttc atc atc tgc	288
Val Gly Pro Gly Lys Pro Leu Asp Pro Ala Arg His Phe Ile Ile Cys	
85 90 95	
tcg aac gtg atc ggc ggc tgc atg ggc tcg acg ggc ccg gcc tcg atc	336
Ser Asn Val Ile Gly Gly Cys Met Gly Ser Thr Gly Pro Ala Ser Ile	
100 105 110	
aat ccg gcc acg ggc aag acc tat ggc ctg tcg ttc cca gtc atc acc	384
Asn Pro Ala Thr Gly Lys Thr Tyr Gly Leu Ser Phe Pro Val Ile Thr	
115 120 125	
atc gcc gat atg gtg cgg gcc cag gcc atg ctg gtc tct gcg ctc ggg	432
Ile Ala Asp Met Val Arg Ala Gln Ala Met Leu Val Ser Ala Leu Gly	
130 135 140	
gtc gag acc ctg ttc gcc gtc gtc ggc ggc tcg atg ggc ggc atg cag	480
Val Glu Thr Leu Phe Ala Val Val Gly Gly Ser Met Gly Gly Met Gln	
145 150 155 160	
gtc cag caa tgg gcc gtg gac tat ccc gag cgg atg ttc agc gcc gtg	528
Val Gln Gln Trp Ala Val Asp Tyr Pro Glu Arg Met Phe Ser Ala Val	
165 170 175	
gtg ctg gcc tcg gcc tcg cgc cac tcg gcc cag aac atc gcg ttc cac	576
Val Leu Ala Ser Ala Ser Arg His Ser Ala Gln Asn Ile Ala Phe His	
180 185 190	
gag gtg ggc cgc cag gcg atc atg gcc gat ccc gac tgg cgc ggc ggc	624
Glu Val Gly Arg Gln Ala Ile Met Ala Asp Pro Asp Trp Arg Gly Gly	
195 200 205	
gcc tat gcc gag cac ggc gtg cgg ccc gag aag ggc ctg gcc gtg gcg	672

Ala	Tyr	Ala	Glu	His	Gly	Val	Arg	Pro	Glu	Lys	Gly	Leu	Ala	Val	Ala		
210						215					220						
cgg	atg	gcc	gcg	cac	atc	acc	tat	ctg	tcc	gag	ccc	gcc	ctg	cag	cgg	720	
Arg	Met	Ala	Ala	His	Ile	Thr	Tyr	Leu	Ser	Glu	Pro	Ala	Leu	Gln	Arg		
225					230					235					240		
aag	ttc	ggc	cgc	gag	cta	cag	cgc	gac	ggc	ctc	tcc	tgg	ggc	ttt	gac	768	
Lys	Phe	Gly	Arg	Glu	Leu	Gln	Arg	Asp	Gly	Leu	Ser	Trp	Gly	Phe	Asp		
				245					250					255			
gcc	gac	ttc	cag	gtc	gag	agc	tat	cta	cgc	cac	cag	ggg	tcc	agc	ttc	816	
Ala	Asp	Phe	Gln	Val	Glu	Ser	Tyr	Leu	Arg	His	Gln	Gly	Ser	Ser	Phe		
			260					265					270				
gtc	gac	cgg	ttc	gac	gcc	aac	agc	tat	ctc	tac	atc	acc	cgg	gcc	atg	864	
Val	Asp	Arg	Phe	Asp	Ala	Asn	Ser	Tyr	Leu	Tyr	Ile	Thr	Arg	Ala	Met		
			275				280					285					
gac	tat	ttc	gac	atc	gcc	gcc	agc	cat	ggc	ggg	gtg	ctg	gcc	aag	gcg	912	
Asp	Tyr	Phe	Asp	Ile	Ala	Ala	Ser	His	Gly	Gly	Val	Leu	Ala	Lys	Ala		
			290			295					300						
ttc	acc	cga	gcg	cgg	aat	gtg	cgc	ttc	tgc	gtg	ctg	agc	ttc	tcc	agc	960	
Phe	Thr	Arg	Ala	Arg	Asn	Val	Arg	Phe	Cys	Val	Leu	Ser	Phe	Ser	Ser		
305					310					315					320		
gac	tgg	ctc	tat	ccg	acc	gcc	gag	aac	cgc	cac	ctg	gtc	cgc	gcc	ctg	1008	
Asp	Trp	Leu	Tyr	Pro	Thr	Ala	Glu	Asn	Arg	His	Leu	Val	Arg	Ala	Leu		
				325					330					335			
acc	gcc	gcc	ggg	gcc	cgc	gcg	gcc	ttc	gcc	gag	atc	gag	agc	gac	aag	1056	
Thr	Ala	Ala	Gly	Ala	Arg	Ala	Ala	Phe	Ala	Glu	Ile	Glu	Ser	Asp	Lys		
			340					345					350				
ggc	cat	gac	gcc	ttc	ctg	ctg	gac	gag	ccg	gtg	atg	gac	gcc	gcg	ctg	1104	
Gly	His	Asp	Ala	Phe	Leu	Leu	Asp	Glu	Pro	Val	Met	Asp	Ala	Ala	Leu		
			355				360					365					
gaa	ggc	ttc	ctg	gcc	tcg	gcc	gaa	cgc	gat	cgg	ggg	ctg	gtt			1146	
Glu	Gly	Phe	Leu	Ala	Ser	Ala	Glu	Arg	Asp	Arg	Gly	Leu	Val				
			370			375					380						
tga																1149	

<210> 10

<211> 382

<212> PRT

<213> Caulobacter crescentus

<400> 10

Met	Ala	Ala	Leu	Asp	Pro	Ile	Thr	Pro	Ala	Gly	Gly	Gly	Thr	Trp	Arg		
1				5					10					15			
Phe	Pro	Ala	Asn	Glu	Pro	Leu	Arg	Leu	Asp	Ser	Gly	Gly	Val	Ile	Glu		
			20					25					30				
Gly	Leu	Glu	Ile	Ala	Tyr	Gln	Thr	Tyr	Gly	Gln	Leu	Asn	Ala	Asp	Lys		
			35				40					45					
Ser	Asn	Ala	Val	Leu	Ile	Cys	His	Ala	Leu	Thr	Gly	Asp	Gln	His	Val		
	50					55					60						

Ala Ser Pro His Pro Thr Thr Gly Lys Pro Gly Trp Trp Gln Arg Leu
 65 70 75 80
 Val Gly Pro Gly Lys Pro Leu Asp Pro Ala Arg His Phe Ile Ile Cys
 85 90 95
 Ser Asn Val Ile Gly Gly Cys Met Gly Ser Thr Gly Pro Ala Ser Ile
 100 105 110
 Asn Pro Ala Thr Gly Lys Thr Tyr Gly Leu Ser Phe Pro Val Ile Thr
 115 120 125
 Ile Ala Asp Met Val Arg Ala Gln Ala Met Leu Val Ser Ala Leu Gly
 130 135 140
 Val Glu Thr Leu Phe Ala Val Val Gly Gly Ser Met Gly Gly Met Gln
 145 150 155 160
 Val Gln Gln Trp Ala Val Asp Tyr Pro Glu Arg Met Phe Ser Ala Val
 165 170 175
 Val Leu Ala Ser Ala Ser Arg His Ser Ala Gln Asn Ile Ala Phe His
 180 185 190
 Glu Val Gly Arg Gln Ala Ile Met Ala Asp Pro Asp Trp Arg Gly Gly
 195 200 205
 Ala Tyr Ala Glu His Gly Val Arg Pro Glu Lys Gly Leu Ala Val Ala
 210 215 220
 Arg Met Ala Ala His Ile Thr Tyr Leu Ser Glu Pro Ala Leu Gln Arg
 225 230 235 240
 Lys Phe Gly Arg Glu Leu Gln Arg Asp Gly Leu Ser Trp Gly Phe Asp
 245 250 255
 Ala Asp Phe Gln Val Glu Ser Tyr Leu Arg His Gln Gly Ser Ser Phe
 260 265 270
 Val Asp Arg Phe Asp Ala Asn Ser Tyr Leu Tyr Ile Thr Arg Ala Met
 275 280 285
 Asp Tyr Phe Asp Ile Ala Ala Ser His Gly Gly Val Leu Ala Lys Ala
 290 295 300
 Phe Thr Arg Ala Arg Asn Val Arg Phe Cys Val Leu Ser Phe Ser Ser
 305 310 315 320
 Asp Trp Leu Tyr Pro Thr Ala Glu Asn Arg His Leu Val Arg Ala Leu
 325 330 335
 Thr Ala Ala Gly Ala Arg Ala Ala Phe Ala Glu Ile Glu Ser Asp Lys
 340 345 350
 Gly His Asp Ala Phe Leu Leu Asp Glu Pro Val Met Asp Ala Ala Leu
 355 360 365
 Glu Gly Phe Leu Ala Ser Ala Glu Arg Asp Arg Gly Leu Val
 370 375 380

<211> 1140
 <212> DNA
 <213> Neisseria gonorrhoeae

<220>
 <221> CDS
 <222> (1)..(1137)
 <223> RNG00132

<400> 11
 atg agt caa aat acc tcg gtg ggc att gta acg ccc caa aaa att ccg 48
 Met Ser Gln Asn Thr Ser Val Gly Ile Val Thr Pro Gln Lys Ile Pro
 1 5 10 15
 ttt gaa atg ccg ctg gtt ttg gaa aac ggt aaa act ttg ccg cgt ttc 96
 Phe Glu Met Pro Leu Val Leu Glu Asn Gly Lys Thr Leu Pro Arg Phe
 20 25 30
 gat ctg atg att gaa acc tac ggc gag ctg aat gct gaa aaa aac aat 144
 Asp Leu Met Ile Glu Thr Tyr Gly Glu Leu Asn Ala Glu Lys Asn Asn
 35 40 45
 gcg gtt tta atc tgc cac gcg ctg tcg ggc aac cat cac gtt gcg ggc 192
 Ala Val Leu Ile Cys His Ala Leu Ser Gly Asn His His Val Ala Gly
 50 55 60
 agg cat tcg gcg gag gat aaa tat acg ggc tgg tgg gac aat atg gtc 240
 Arg His Ser Ala Glu Asp Lys Tyr Thr Gly Trp Trp Asp Asn Met Val
 65 70 75 80
 ggt ccc gga aaa ccg att gat acg gaa cgt ttt ttc gtg gtc ggg ttg 288
 Gly Pro Gly Lys Pro Ile Asp Thr Glu Arg Phe Phe Val Val Gly Leu
 85 90 95
 aac aat ctg ggc ggc tgc gac ggc agc agc ggg cct ttg tcg atc aat 336
 Asn Asn Leu Gly Gly Cys Asp Gly Ser Ser Gly Pro Leu Ser Ile Asn
 100 105 110
 cct gaa acg ggc agg gaa tac ggc gcg gat ttt ccg atg gtt acg gtg 384
 Pro Glu Thr Gly Arg Glu Tyr Gly Ala Asp Phe Pro Met Val Thr Val
 115 120 125
 aag gac tgg gta aaa tca caa gcc gcg ctt gcc gat tat ctc ggc atc 432
 Lys Asp Trp Val Lys Ser Gln Ala Ala Leu Ala Asp Tyr Leu Gly Ile
 130 135 140
 gaa caa tgg gcg gcg gtt gtc ggc ggc agc ttg ggc ggc atg cag gct 480
 Glu Gln Trp Ala Ala Val Val Gly Gly Ser Leu Gly Gly Met Gln Ala
 145 150 155 160
 ttg cag tgg gcg att tcc tat ccc gaa cgt gtg cgc cac gcc ttg gtg 528
 Leu Gln Trp Ala Ile Ser Tyr Pro Glu Arg Val Arg His Ala Leu Val
 165 170 175
 att gcg tct gcg ccg aaa ctg tcc gcg caa aat atc gcg ttt aat gat 576
 Ile Ala Ser Ala Pro Lys Leu Ser Ala Gln Asn Ile Ala Phe Asn Asp
 180 185 190
 gta gca cgt cag gcg att ttg acc gac ccc gat ttc aat gaa gga cat 624
 Val Ala Arg Gln Ala Ile Leu Thr Asp Pro Asp Phe Asn Glu Gly His
 195 200 205
 tac cgc agc cac aac acc gtt ccc gcg cgc ggt ttg cgg att gcc cgt 672

Tyr Arg Ser His Asn Thr Val Pro Ala Arg Gly Leu Arg Ile Ala Arg
 210 215 220
 atg atg gga cac att acg tat ctt gcc gaa gac ggt ttg ggc aaa aaa 720
 Met Met Gly His Ile Thr Tyr Leu Ala Glu Asp Gly Leu Gly Lys Lys
 225 230 235 240
 ttc gga cgc gat ttg cgt tcc aac ggc tat caa tac ggc tat agc gtt 768
 Phe Gly Arg Asp Leu Arg Ser Asn Gly Tyr Gln Tyr Gly Tyr Ser Val
 245 250 255
 gaa ttt gaa gta gaa tcc tat ctc cgc tat caa ggc gac aaa ttc gtc 816
 Glu Phe Glu Val Glu Ser Tyr Leu Arg Tyr Gln Gly Asp Lys Phe Val
 260 265 270
 ggg cgg ttt gat gct aat aca tat ttg ctg atg acc aaa gct ttg gac 864
 Gly Arg Phe Asp Ala Asn Thr Tyr Leu Leu Met Thr Lys Ala Leu Asp
 275 280 285
 tat ttc gat ccg gcg gcg gat ttc ggc aac agc ctg acc cgc gcc gtg 912
 Tyr Phe Asp Pro Ala Ala Asp Phe Gly Asn Ser Leu Thr Arg Ala Val
 290 295 300
 cag gat gtg cag gca aaa ttc ttt gtc gcc agc ttc agc acc gac tgg 960
 Gln Asp Val Gln Ala Lys Phe Phe Val Ala Ser Phe Ser Thr Asp Trp
 305 310 315 320
 cgt ttc gcg ccc gaa cgt tcg cac gaa ctg gtc aag gca ctg att gcc 1008
 Arg Phe Ala Pro Glu Arg Ser His Glu Leu Val Lys Ala Leu Ile Ala
 325 330 335
 gcc caa aaa tcc gtg cag tat atc gaa gtc aag tcc gca cac ggg cac 1056
 Ala Gln Lys Ser Val Gln Tyr Ile Glu Val Lys Ser Ala His Gly His
 340 345 350
 gat gcc ttt tta atg gaa gac gaa gcc tat atg cgc gcc gta acg gct 1104
 Asp Ala Phe Leu Met Glu Asp Glu Ala Tyr Met Arg Ala Val Thr Ala
 355 360 365
 tat atg aac aat gtt gac aag gat tgc cga tta tga 1140
 Tyr Met Asn Asn Val Asp Lys Asp Cys Arg Leu
 370 375

<210> 12

<211> 379

<212> PRT

<213> Neisseria gonorrhoeae

<400> 12

Met Ser Gln Asn Thr Ser Val Gly Ile Val Thr Pro Gln Lys Ile Pro
 1 5 10 15
 Phe Glu Met Pro Leu Val Leu Glu Asn Gly Lys Thr Leu Pro Arg Phe
 20 25 30
 Asp Leu Met Ile Glu Thr Tyr Gly Glu Leu Asn Ala Glu Lys Asn Asn
 35 40 45
 Ala Val Leu Ile Cys His Ala Leu Ser Gly Asn His His Val Ala Gly
 50 55 60

Arg His Ser Ala Glu Asp Lys Tyr Thr Gly Trp Trp Asp Asn Met Val
 65 70 75 80
 Gly Pro Gly Lys Pro Ile Asp Thr Glu Arg Phe Phe Val Val Gly Leu
 85 90 95
 Asn Asn Leu Gly Gly Cys Asp Gly Ser Ser Gly Pro Leu Ser Ile Asn
 100 105 110
 Pro Glu Thr Gly Arg Glu Tyr Gly Ala Asp Phe Pro Met Val Thr Val
 115 120 125
 Lys Asp Trp Val Lys Ser Gln Ala Ala Leu Ala Asp Tyr Leu Gly Ile
 130 135 140
 Glu Gln Trp Ala Ala Val Val Gly Gly Ser Leu Gly Gly Met Gln Ala
 145 150 155 160
 Leu Gln Trp Ala Ile Ser Tyr Pro Glu Arg Val Arg His Ala Leu Val
 165 170 175
 Ile Ala Ser Ala Pro Lys Leu Ser Ala Gln Asn Ile Ala Phe Asn Asp
 180 185 190
 Val Ala Arg Gln Ala Ile Leu Thr Asp Pro Asp Phe Asn Glu Gly His
 195 200 205
 Tyr Arg Ser His Asn Thr Val Pro Ala Arg Gly Leu Arg Ile Ala Arg
 210 215 220
 Met Met Gly His Ile Thr Tyr Leu Ala Glu Asp Gly Leu Gly Lys Lys
 225 230 235 240
 Phe Gly Arg Asp Leu Arg Ser Asn Gly Tyr Gln Tyr Gly Tyr Ser Val
 245 250 255
 Glu Phe Glu Val Glu Ser Tyr Leu Arg Tyr Gln Gly Asp Lys Phe Val
 260 265 270
 Gly Arg Phe Asp Ala Asn Thr Tyr Leu Leu Met Thr Lys Ala Leu Asp
 275 280 285
 Tyr Phe Asp Pro Ala Ala Asp Phe Gly Asn Ser Leu Thr Arg Ala Val
 290 295 300
 Gln Asp Val Gln Ala Lys Phe Phe Val Ala Ser Phe Ser Thr Asp Trp
 305 310 315 320
 Arg Phe Ala Pro Glu Arg Ser His Glu Leu Val Lys Ala Leu Ile Ala
 325 330 335
 Ala Gln Lys Ser Val Gln Tyr Ile Glu Val Lys Ser Ala His Gly His
 340 345 350
 Asp Ala Phe Leu Met Glu Asp Glu Ala Tyr Met Arg Ala Val Thr Ala
 355 360 365
 Tyr Met Asn Asn Val Asp Lys Asp Cys Arg Leu
 370 375

<210> 13

<211> 1140

<212> DNA

<213> *Neisseria meningitidis* ser. A

<220>

<221> CDS

<222> (1)..(1137)

<223> RNM00815

<400> 13

```

atg agt caa aat gcc tcg gtg ggc att gta acg ccc caa aaa att ccg      48
Met Ser Gln Asn Ala Ser Val Gly Ile Val Thr Pro Gln Lys Ile Pro
1           5           10           15

ttt gaa atg ccg ctg gtt ttg gaa aac ggt aaa act ttg ccg cgt ttc      96
Phe Glu Met Pro Leu Val Leu Glu Asn Gly Lys Thr Leu Pro Arg Phe
           20           25           30

gat ctg atg att gaa acc tac ggc gag ctg aat gcc gaa aaa aat aat      144
Asp Leu Met Ile Glu Thr Tyr Gly Glu Leu Asn Ala Glu Lys Asn Asn
           35           40           45

gcg gtt tta atc tgt cat gcg ctg tca ggc aac cat cat gtt gcg ggc      192
Ala Val Leu Ile Cys His Ala Leu Ser Gly Asn His His Val Ala Gly
           50           55           60

agg cat tcg gcg gag gat aaa tat acg ggc tgg tgg gac aat atg gta      240
Arg His Ser Ala Glu Asp Lys Tyr Thr Gly Trp Trp Asp Asn Met Val
           65           70           75           80

gga ccc ggc aaa ccg att gat aca gaa cgt ttt ttc gtg gtc ggt ttg      288
Gly Pro Gly Lys Pro Ile Asp Thr Glu Arg Phe Phe Val Val Gly Leu
           85           90           95

aac aat ctg ggc ggc tgc gac ggc agc agc gga cct ttg tcg atc aat      336
Asn Asn Leu Gly Gly Cys Asp Gly Ser Ser Gly Pro Leu Ser Ile Asn
           100           105           110

cct gaa acg ggc agg gaa tac ggc gcg gat ttt ccg gtg gtt acg gtg      384
Pro Glu Thr Gly Arg Glu Tyr Gly Ala Asp Phe Pro Val Val Thr Val
           115           120           125

aag gac tgg gta aaa tcc caa gcc gcg ctt acc gat tat ctc ggc atc      432
Lys Asp Trp Val Lys Ser Gln Ala Ala Leu Thr Asp Tyr Leu Gly Ile
           130           135           140

ggg caa tgg gcg gcg gtt gtc ggc ggc agc ttg ggc ggt atg cag gct      480
Gly Gln Trp Ala Ala Val Val Gly Gly Ser Leu Gly Gly Met Gln Ala
           145           150           155           160

ttg cag tgg acg att tcc tat ccc gag cgc gtg cgc cat gcc tta gtg      528
Leu Gln Trp Thr Ile Ser Tyr Pro Glu Arg Val Arg His Ala Leu Val
           165           170           175

att gcg tcc gcg ccg aaa ctg tcc acg caa aat atc gcg ttt aat gat      576
Ile Ala Ser Ala Pro Lys Leu Ser Thr Gln Asn Ile Ala Phe Asn Asp
           180           185           190

gta gca cgt cag gcg att ttg acc gat ccc gat ttc aac gaa gga cat      624
Val Ala Arg Gln Ala Ile Leu Thr Asp Pro Asp Phe Asn Glu Gly His
           195           200           205

tac cgc agc cgc aac acc gtt ccc gct cgg ggc ttg cgg att gcc cgc      672
Tyr Arg Ser Arg Asn Thr Val Pro Ala Arg Gly Leu Arg Ile Ala Arg

```

210

215

220

atg atg ggg cac atc acc tat ctt gcc gaa gac ggt ttg ggc aaa aaa 720
 Met Met Gly His Ile Thr Tyr Leu Ala Glu Asp Gly Leu Gly Lys Lys
 225 230 235 240

ttc gga cgc gat ttg cgt tcc aac ggc tat caa tac ggc tat ggc gtt 768
 Phe Gly Arg Asp Leu Arg Ser Asn Gly Tyr Gln Tyr Gly Tyr Gly Val
 245 250 255

gaa ttt gaa gta gaa tcc tat ctg cgc tat caa ggc gat aaa ttc gtc 816
 Glu Phe Glu Val Glu Ser Tyr Leu Arg Tyr Gln Gly Asp Lys Phe Val
 260 265 270

ggg cgg ttt gat gcc aac acc tat ttg ctg atg acc aag gct ttg gac 864
 Gly Arg Phe Asp Ala Asn Thr Tyr Leu Leu Met Thr Lys Ala Leu Asp
 275 280 285

tat ttc gat ccg gcg gcg gat ttc ggc aac agc ctg acc cgc gcc gtg 912
 Tyr Phe Asp Pro Ala Ala Asp Phe Gly Asn Ser Leu Thr Arg Ala Val
 290 295 300

cag gat gtt cag gca aaa ttc ttt gtc gcc agc ttc agc acc gat tgg 960
 Gln Asp Val Gln Ala Lys Phe Phe Val Ala Ser Phe Ser Thr Asp Trp
 305 310 315 320

cgt ttc gcg ccc gaa cgt tcg cac gaa ctg gtc aag gcc ctg att gcc 1008
 Arg Phe Ala Pro Glu Arg Ser His Glu Leu Val Lys Ala Leu Ile Ala
 325 330 335

gcc caa aaa tcc gtg cag tat atc gaa gtc aaa tcc gca cac ggg cac 1056
 Ala Gln Lys Ser Val Gln Tyr Ile Glu Val Lys Ser Ala His Gly His
 340 345 350

gat gcc ttt tta atg gaa gac gaa gcc tat atg cgt gcg gtc gcc gcc 1104
 Asp Ala Phe Leu Met Glu Asp Glu Ala Tyr Met Arg Ala Val Ala Ala
 355 360 365

tat atg aac aac gtt tat aag gaa tgt cag caa tga 1140
 Tyr Met Asn Asn Val Tyr Lys Glu Cys Gln Gln
 370 375

<210> 14

<211> 379

<212> PRT

<213> Neisseria meningitidis ser. A

<400> 14

Met Ser Gln Asn Ala Ser Val Gly Ile Val Thr Pro Gln Lys Ile Pro
 1 5 10 15

Phe Glu Met Pro Leu Val Leu Glu Asn Gly Lys Thr Leu Pro Arg Phe
 20 25 30

Asp Leu Met Ile Glu Thr Tyr Gly Glu Leu Asn Ala Glu Lys Asn Asn
 35 40 45

Ala Val Leu Ile Cys His Ala Leu Ser Gly Asn His His Val Ala Gly
 50 55 60

Arg His Ser Ala Glu Asp Lys Tyr Thr Gly Trp Trp Asp Asn Met Val
 65 70 75 80

Gly Pro Gly Lys Pro Ile Asp Thr Glu Arg Phe Phe Val Val Gly Leu
85 90 95

Asn Asn Leu Gly Gly Cys Asp Gly Ser Ser Gly Pro Leu Ser Ile Asn
100 105 110

Pro Glu Thr Gly Arg Glu Tyr Gly Ala Asp Phe Pro Val Val Thr Val
115 120 125

Lys Asp Trp Val Lys Ser Gln Ala Ala Leu Thr Asp Tyr Leu Gly Ile
130 135 140

Gly Gln Trp Ala Ala Val Val Gly Gly Ser Leu Gly Gly Met Gln Ala
145 150 155 160

Leu Gln Trp Thr Ile Ser Tyr Pro Glu Arg Val Arg His Ala Leu Val
165 170 175

Ile Ala Ser Ala Pro Lys Leu Ser Thr Gln Asn Ile Ala Phe Asn Asp
180 185 190

Val Ala Arg Gln Ala Ile Leu Thr Asp Pro Asp Phe Asn Glu Gly His
195 200 205

Tyr Arg Ser Arg Asn Thr Val Pro Ala Arg Gly Leu Arg Ile Ala Arg
210 215 220

Met Met Gly His Ile Thr Tyr Leu Ala Glu Asp Gly Leu Gly Lys Lys
225 230 235 240

Phe Gly Arg Asp Leu Arg Ser Asn Gly Tyr Gln Tyr Gly Tyr Gly Val
245 250 255

Glu Phe Glu Val Glu Ser Tyr Leu Arg Tyr Gln Gly Asp Lys Phe Val
260 265 270

Gly Arg Phe Asp Ala Asn Thr Tyr Leu Leu Met Thr Lys Ala Leu Asp
275 280 285

Tyr Phe Asp Pro Ala Ala Asp Phe Gly Asn Ser Leu Thr Arg Ala Val
290 295 300

Gln Asp Val Gln Ala Lys Phe Phe Val Ala Ser Phe Ser Thr Asp Trp
305 310 315 320

Arg Phe Ala Pro Glu Arg Ser His Glu Leu Val Lys Ala Leu Ile Ala
325 330 335

Ala Gln Lys Ser Val Gln Tyr Ile Glu Val Lys Ser Ala His Gly His
340 345 350

Asp Ala Phe Leu Met Glu Asp Glu Ala Tyr Met Arg Ala Val Ala Ala
355 360 365

Tyr Met Asn Asn Val Tyr Lys Glu Cys Gln Gln
370 375

```
<210> 15
<211> 1140
<212> DNA
<213> Pseudomonas fluorescens
```

<220>

<221> CDS

<222> (1) .. (1137)

<223> RPU01633

<400> 15

```

atg cca gct gcc ttt ccc ccc gat tct gtt ggt ctg gtg acg ccg caa      48
Met Pro Ala Ala Phe Pro Pro Asp Ser Val Gly Leu Val Thr Pro Gln
  1              5              10              15

acg gcg cac ttc agc gaa ccg ctg gcc ctg gcc tgc ggc cgt tcg ctg      96
Thr Ala His Phe Ser Glu Pro Leu Ala Leu Ala Cys Gly Arg Ser Leu
              20              25              30

gcc gat tat gac ctg atc tac gaa acc tac ggc acg ctg aac gcg caa      144
Ala Asp Tyr Asp Leu Ile Tyr Glu Thr Tyr Gly Thr Leu Asn Ala Gln
              35              40              45

gcg agc aac gcc gtg ctg atc tgc cac gcc ttg tcc ggc cac cac cat      192
Ala Ser Asn Ala Val Leu Ile Cys His Ala Leu Ser Gly His His His
              50              55              60

gct gcg ggt tat cac agc gtc gac gac cgc aag ccc ggt tgg tgg gac      240
Ala Ala Gly Tyr His Ser Val Asp Asp Arg Lys Pro Gly Trp Trp Asp
  65              70              75              80

agc tgc atc ggc ccc ggc aaa ccg atc gac acc aac aag ttc ttc gtg      288
Ser Cys Ile Gly Pro Gly Lys Pro Ile Asp Thr Asn Lys Phe Phe Val
              85              90              95

gtc agc ctg aac aac ctc ggc ggt tgc aat ggt tct acc ggc ccg agc      336
Val Ser Leu Asn Asn Leu Gly Gly Cys Asn Gly Ser Thr Gly Pro Ser
              100              105              110

agc ctc aat ccg gaa acc ggc aag ccg ttc ggc gcc gac ttc ccg gtg      384
Ser Leu Asn Pro Glu Thr Gly Lys Pro Phe Gly Ala Asp Phe Pro Val
              115              120              125

ctg acc gtg gaa gac tgg gtg cac agc cag gca cgc ctg gcc gac ctg      432
Leu Thr Val Glu Asp Trp Val His Ser Gln Ala Arg Leu Ala Asp Leu
              130              135              140

ctc ggc atc ggc cag tgg gcg gcg gtg atc ggc ggc agc ctg ggc gcc      480
Leu Gly Ile Gly Gln Trp Ala Ala Val Ile Gly Gly Ser Leu Gly Gly
  145              150              155              160

atg cag gcg ctg caa tgg acc atc acc tat ccg gat cgc gtt cgc cac      528
Met Gln Ala Leu Gln Trp Thr Ile Thr Tyr Pro Asp Arg Val Arg His
              165              170              175

tgc ctg gcc atc gcc tcg gcc ccc aag ctg tcg gcg cag aac atc gcc      576
Cys Leu Ala Ile Ala Ser Ala Pro Lys Leu Ser Ala Gln Asn Ile Ala
              180              185              190

ttc aac gaa gtg gcg cgc cag gcg atc ctc act gac ccg gaa ttc cac      624
Phe Asn Glu Val Ala Arg Gln Ala Ile Leu Thr Asp Pro Glu Phe His
              195              200              205

ggc gcc tcg ttc cag gaa cac gcc gtg atc ccc aag cgc gcc ctg atg      672
Gly Gly Ser Phe Gln Glu His Gly Val Ile Pro Lys Arg Gly Leu Met
              210              215              220

```

23/92

ctg gcg cgg atg gtg ggg cac atc acc tac ctg tcc gac gac tcc atg 720
 Leu Ala Arg Met Val Gly His Ile Thr Tyr Leu Ser Asp Asp Ser Met
 225 230 235 240

ggt gag aaa ttc ggc cgt ggc ctg aag agc gaa aag ctc aac tac gac 768
 Gly Glu Lys Phe Gly Arg Gly Leu Lys Ser Glu Lys Leu Asn Tyr Asp
 245 250 255

ttc cac agc gtc gag ttc cag gtc gaa agc tac ctg cgc tat cag ggc 816
 Phe His Ser Val Glu Phe Gln Val Glu Ser Tyr Leu Arg Tyr Gln Gly
 260 265 270

gaa gag ttc tcc ggg cgc ttc gat gcc aac acc tat ctg ttg atg acc 864
 Glu Glu Phe Ser Gly Arg Phe Asp Ala Asn Thr Tyr Leu Leu Met Thr
 275 280 285

aag gcg ctg gac tac ttc gat ccg gcg gcg aac ttc aac gat aac ctg 912
 Lys Ala Leu Asp Tyr Phe Asp Pro Ala Ala Asn Phe Asn Asp Asn Leu
 290 295 300

gcg aaa acc ttc gaa ggt gca aaa gcc aag ttc tgc gtg atg tgc ttc 960
 Ala Lys Thr Phe Glu Gly Ala Lys Ala Lys Phe Cys Val Met Ser Phe
 305 310 315 320

acc acc gac tgg cgc ttc tcc ccg gcc cgc tgc cga gaa ctg gtg gat 1008
 Thr Thr Asp Trp Arg Phe Ser Pro Ala Arg Ser Arg Glu Leu Val Asp
 325 330 335

gcg ctg atg gcg gcg cgc aaa gac gtc agc tac ctg gaa atc gac gcg 1056
 Ala Leu Met Ala Ala Arg Lys Asp Val Ser Tyr Leu Glu Ile Asp Ala
 340 345 350

ccc cag ggc cac gac gcc ttc ctg att ccg atc ccg cgc tac ttg cag 1104
 Pro Gln Gly His Asp Ala Phe Leu Ile Pro Ile Pro Arg Tyr Leu Gln
 355 360 365

gcg ttc ggc aat tac atg aac cgc att acg ttg tga 1140
 Ala Phe Gly Asn Tyr Met Asn Arg Ile Thr Leu
 370 375

<210> 16
 <211> 379
 <212> PRT
 <213> Pseudomonas fluorescens

<400> 16
 Met Pro Ala Ala Phe Pro Pro Asp Ser Val Gly Leu Val Thr Pro Gln
 1 5 10 15
 Thr Ala His Phe Ser Glu Pro Leu Ala Leu Ala Cys Gly Arg Ser Leu
 20 25 30
 Ala Asp Tyr Asp Leu Ile Tyr Glu Thr Tyr Gly Thr Leu Asn Ala Gln
 35 40 45
 Ala Ser Asn Ala Val Leu Ile Cys His Ala Leu Ser Gly His His His
 50 55 60
 Ala Ala Gly Tyr His Ser Val Asp Asp Arg Lys Pro Gly Trp Trp Asp
 65 70 75 80
 Ser Cys Ile Gly Pro Gly Lys Pro Ile Asp Thr Asn Lys Phe Phe Val

85 90 95
 Val Ser Leu Asn Asn Leu Gly Gly Cys Asn Gly Ser Thr Gly Pro Ser
 100 105 110
 Ser Leu Asn Pro Glu Thr Gly Lys Pro Phe Gly Ala Asp Phe Pro Val
 115 120 125
 Leu Thr Val Glu Asp Trp Val His Ser Gln Ala Arg Leu Ala Asp Leu
 130 135 140
 Leu Gly Ile Gly Gln Trp Ala Ala Val Ile Gly Gly Ser Leu Gly Gly
 145 150 155 160
 Met Gln Ala Leu Gln Trp Thr Ile Thr Tyr Pro Asp Arg Val Arg His
 165 170 175
 Cys Leu Ala Ile Ala Ser Ala Pro Lys Leu Ser Ala Gln Asn Ile Ala
 180 185 190
 Phe Asn Glu Val Ala Arg Gln Ala Ile Leu Thr Asp Pro Glu Phe His
 195 200 205
 Gly Gly Ser Phe Gln Glu His Gly Val Ile Pro Lys Arg Gly Leu Met
 210 215 220
 Leu Ala Arg Met Val Gly His Ile Thr Tyr Leu Ser Asp Asp Ser Met
 225 230 235 240
 Gly Glu Lys Phe Gly Arg Gly Leu Lys Ser Glu Lys Leu Asn Tyr Asp
 245 250 255
 Phe His Ser Val Glu Phe Gln Val Glu Ser Tyr Leu Arg Tyr Gln Gly
 260 265 270
 Glu Glu Phe Ser Gly Arg Phe Asp Ala Asn Thr Tyr Leu Leu Met Thr
 275 280 285
 Lys Ala Leu Asp Tyr Phe Asp Pro Ala Ala Asn Phe Asn Asp Asn Leu
 290 295 300
 Ala Lys Thr Phe Glu Gly Ala Lys Ala Lys Phe Cys Val Met Ser Phe
 305 310 315 320
 Thr Thr Asp Trp Arg Phe Ser Pro Ala Arg Ser Arg Glu Leu Val Asp
 325 330 335
 Ala Leu Met Ala Ala Arg Lys Asp Val Ser Tyr Leu Glu Ile Asp Ala
 340 345 350
 Pro Gln Gly His Asp Ala Phe Leu Ile Pro Ile Pro Arg Tyr Leu Gln
 355 360 365
 Ala Phe Gly Asn Tyr Met Asn Arg Ile Thr Leu
 370 375

<210> 17

<211> 1140

<212> DNA

<213> Pseudomonas aeruginosa

<220>

<221> CDS

<222> (1) .. (1137)

<223> RPA04460

<400> 17

```

atg ccc aca gtc ttc ccc gac gac tcc gtc ggt ctg gtc tcc ccc cag 48
Met Pro Thr Val Phe Pro Asp Asp Ser Val Gly Leu Val Ser Pro Gln
1 5 10 15

acg ctg cac ttc aac gaa ccg ctc gag ctg acc agc ggc aag tcc ctg 96
Thr Leu His Phe Asn Glu Pro Leu Glu Leu Thr Ser Gly Lys Ser Leu
20 25 30

gcc gag tac gac ctg gtg atc gaa acc tac ggc gag ctg aat gcc acg 144
Ala Glu Tyr Asp Leu Val Ile Glu Thr Tyr Gly Glu Leu Asn Ala Thr
35 40 45

cag agc aac gcg gtg ctg atc tgc cac gcc ctc tcc ggc cac cac cac 192
Gln Ser Asn Ala Val Leu Ile Cys His Ala Leu Ser Gly His His His
50 55 60

gcc gcc ggc tac cac agc gtc gac gag cgc aag ccg ggc tgg tgg gac 240
Ala Ala Gly Tyr His Ser Val Asp Glu Arg Lys Pro Gly Trp Trp Asp
65 70 75 80

agc tgc atc ggt ccg ggc aag ccg atc gac acc cgc aag ttc ttc gtc 288
Ser Cys Ile Gly Pro Gly Lys Pro Ile Asp Thr Arg Lys Phe Phe Val
85 90 95

gtc gcc ctc aac aac ctc ggc ggt tgc aac gga tcc agc ggc ccc gcc 336
Val Ala Leu Asn Asn Leu Gly Gly Cys Asn Gly Ser Ser Gly Pro Ala
100 105 110

agc atc aat ccg gcg acc ggc aag gtc tac ggc gcg gac ttc ccg atg 384
Ser Ile Asn Pro Ala Thr Gly Lys Val Tyr Gly Ala Asp Phe Pro Met
115 120 125

gtt acg gtg gaa gac tgg gtg cat agc cag gcg cgc ctg gca gac cgc 432
Val Thr Val Glu Asp Trp Val His Ser Gln Ala Arg Leu Ala Asp Arg
130 135 140

ctc ggc atc cgc cag tgg gcc gcg gtg gtc ggc ggc agc ctc ggc ggc 480
Leu Gly Ile Arg Gln Trp Ala Ala Val Val Gly Gly Ser Leu Gly Gly
145 150 155 160

atg cag gcg ctg caa tgg acc atc agc tat ccc gag cgc gtc cgt cac 528
Met Gln Ala Leu Gln Trp Thr Ile Ser Tyr Pro Glu Arg Val Arg His
165 170 175

tgc ctg tgc atc gcc agc gcg ccg aag ctg tcg gcg cag aac atc gcc 576
Cys Leu Cys Ile Ala Ser Ala Pro Lys Leu Ser Ala Gln Asn Ile Ala
180 185 190

ttc aac gaa gtc gcc ccg cag gcg att ctt tcc gac cct gag ttc ctc 624
Phe Asn Glu Val Ala Arg Gln Ala Ile Leu Ser Asp Pro Glu Phe Leu
195 200 205

ggc ggc tac ttc cag gag cag ggc gtg att ccc aag cgc ggc ctc aag 672
Gly Gly Tyr Phe Gln Glu Gln Gly Val Ile Pro Lys Arg Gly Leu Lys
210 215 220

ctg gcg ccg atg gtc ggc cat atc acc tac ctg tcc gac gac gcc atg 720
Leu Ala Arg Met Val Gly His Ile Thr Tyr Leu Ser Asp Asp Ala Met

```

225	230	235	240	
ggc gcc aag ttc ggc cgt gta ctg aag acc gag aag ctc aac tac gac				768
Gly Ala Lys Phe Gly Arg Val Leu Lys Thr Glu Lys Leu Asn Tyr Asp				
245		250	255	
ctg cac agc gtc gag ttc cag gtc gag agt tac ctg cgc tac cag ggc				816
Leu His Ser Val Glu Phe Gln Val Glu Ser Tyr Leu Arg Tyr Gln Gly				
260		265	270	
gag gag ttc tcc acc cgc ttc gac gcc aat acc tac ctg ctg atg acc				864
Glu Glu Phe Ser Thr Arg Phe Asp Ala Asn Thr Tyr Leu Leu Met Thr				
275		280	285	
aag gcg ctg gac tac ttc gac ccc gcc gcc gcc cac ggc gac gac ctg				912
Lys Ala Leu Asp Tyr Phe Asp Pro Ala Ala Ala His Gly Asp Asp Leu				
290		295	300	
gtg cgc acc ctg gag ggc gtc gag gcg gac ttc tgc ctg atg tcc ttc				960
Val Arg Thr Leu Glu Gly Val Glu Ala Asp Phe Cys Leu Met Ser Phe				
305		310	315	320
acc acc gac tgg cgt ttc tcg ccg gcc cgc tcg cgg gaa atc gtc gac				1008
Thr Thr Asp Trp Arg Phe Ser Pro Ala Arg Ser Arg Glu Ile Val Asp				
325		330	335	
gcc ctg atc gcg gcg aaa aag aac gtc agc tac ctg gag atc gac gcc				1056
Ala Leu Ile Ala Ala Lys Lys Asn Val Ser Tyr Leu Glu Ile Asp Ala				
340		345	350	
ccg caa ggc cac gac gcc ttc ctc atg ccg atc ccc cgg tac ctg caa				1104
Pro Gln Gly His Asp Ala Phe Leu Met Pro Ile Pro Arg Tyr Leu Gln				
355		360	365	
gcc ttc agc ggt tac atg aac cgc atc agc gtg tga				1140
Ala Phe Ser Gly Tyr Met Asn Arg Ile Ser Val				
370		375		

<210> 18

<211> 379

<212> PRT

<213> Pseudomonas aeruginosa

<400> 18

Met Pro Thr Val Phe Pro Asp Asp Ser Val Gly Leu Val Ser Pro Gln				
1	5	10	15	
Thr Leu His Phe Asn Glu Pro Leu Glu Leu Thr Ser Gly Lys Ser Leu				
20		25	30	
Ala Glu Tyr Asp Leu Val Ile Glu Thr Tyr Gly Glu Leu Asn Ala Thr				
35		40	45	
Gln Ser Asn Ala Val Leu Ile Cys His Ala Leu Ser Gly His His His				
50		55	60	
Ala Ala Gly Tyr His Ser Val Asp Glu Arg Lys Pro Gly Trp Trp Asp				
65		70	75	80
Ser Cys Ile Gly Pro Gly Lys Pro Ile Asp Thr Arg Lys Phe Phe Val				
85		90	95	

Val Ala Leu Asn Asn Leu Gly Gly Cys Asn Gly Ser Ser Gly Pro Ala
 100 105 110
 Ser Ile Asn Pro Ala Thr Gly Lys Val Tyr Gly Ala Asp Phe Pro Met
 115 120 125
 Val Thr Val Glu Asp Trp Val His Ser Gln Ala Arg Leu Ala Asp Arg
 130 135 140
 Leu Gly Ile Arg Gln Trp Ala Ala Val Val Gly Gly Ser Leu Gly Gly
 145 150 155 160
 Met Gln Ala Leu Gln Trp Thr Ile Ser Tyr Pro Glu Arg Val Arg His
 165 170 175
 Cys Leu Cys Ile Ala Ser Ala Pro Lys Leu Ser Ala Gln Asn Ile Ala
 180 185 190
 Phe Asn Glu Val Ala Arg Gln Ala Ile Leu Ser Asp Pro Glu Phe Leu
 195 200 205
 Gly Gly Tyr Phe Gln Glu Gln Gly Val Ile Pro Lys Arg Gly Leu Lys
 210 215 220
 Leu Ala Arg Met Val Gly His Ile Thr Tyr Leu Ser Asp Asp Ala Met
 225 230 235 240
 Gly Ala Lys Phe Gly Arg Val Leu Lys Thr Glu Lys Leu Asn Tyr Asp
 245 250 255
 Leu His Ser Val Glu Phe Gln Val Glu Ser Tyr Leu Arg Tyr Gln Gly
 260 265 270
 Glu Glu Phe Ser Thr Arg Phe Asp Ala Asn Thr Tyr Leu Leu Met Thr
 275 280 285
 Lys Ala Leu Asp Tyr Phe Asp Pro Ala Ala Ala His Gly Asp Asp Leu
 290 295 300
 Val Arg Thr Leu Glu Gly Val Glu Ala Asp Phe Cys Leu Met Ser Phe
 305 310 315 320
 Thr Thr Asp Trp Arg Phe Ser Pro Ala Arg Ser Arg Glu Ile Val Asp
 325 330 335
 Ala Leu Ile Ala Ala Lys Lys Asn Val Ser Tyr Leu Glu Ile Asp Ala
 340 345 350
 Pro Gln Gly His Asp Ala Phe Leu Met Pro Ile Pro Arg Tyr Leu Gln
 355 360 365
 Ala Phe Ser Gly Tyr Met Asn Arg Ile Ser Val
 370 375

<210> 19

<211> 1146

<212> DNA

<213> Burkholderia cepacia

<220>

<221> CDS

<222> (1) .. (1143)

<223> RBU12675

<400> 19

atg gaa tcg atc ggt atc gtc gct ccc caa aaa atg cat ttc acc gag	48
Met Glu Ser Ile Gly Ile Val Ala Pro Gln Lys Met His Phe Thr Glu	
1 5 10 15	
ccg ctg ccg ttg cag aac ggc agt tcg ctc gcc ggt tac gac ctg atg	96
Pro Leu Pro Leu Gln Asn Gly Ser Ser Leu Ala Gly Tyr Asp Leu Met	
20 25 30	
gtc gag acc tac ggc acg ctc aac gcc gcg cgt agc aac gcg gtg ctg	144
Val Glu Thr Tyr Gly Thr Leu Asn Ala Ala Arg Ser Asn Ala Val Leu	
35 40 45	
gtg tgc cac gcg ctc aac gcg tcg cac cac gtg gcg ggc gtg tat gcc	192
Val Cys His Ala Leu Asn Ala Ser His His Val Ala Gly Val Tyr Ala	
50 55 60	
gac aac ccc agg gac atc ggc tgg tgg gac aac atg gtc ggc ccg ggc	240
Asp Asn Pro Arg Asp Ile Gly Trp Trp Asp Asn Met Val Gly Pro Gly	
65 70 75 80	
aag ccg ctc gac act gac aag ttc ttc gtg atc ggc gtg aac aac ctc	288
Lys Pro Leu Asp Thr Asp Lys Phe Phe Val Ile Gly Val Asn Asn Leu	
85 90 95	
gga tcg tgc ttc ggc tcg act ggg ccg atg agc atc gat ccg tct acc	336
Gly Ser Cys Phe Gly Ser Thr Gly Pro Met Ser Ile Asp Pro Ser Thr	
100 105 110	
ggc aat ccg tac ggc gcg acg ttt ccc gtc gtg acg gtg gaa gac tgg	384
Gly Asn Pro Tyr Gly Ala Thr Phe Pro Val Val Thr Val Glu Asp Trp	
115 120 125	
gtc aac gcc cag gcg cgc gtc gcg gat caa ttc ggc atc acg cgc ttt	432
Val Asn Ala Gln Ala Arg Val Ala Asp Gln Phe Gly Ile Thr Arg Phe	
130 135 140	
gcg gcg gtg atg ggc ggc agc ctc ggc ggc atg cag gcg ctc gcg tgg	480
Ala Ala Val Met Gly Gly Ser Leu Gly Gly Met Gln Ala Leu Ala Trp	
145 150 155 160	
agc atg atg tat ccg gag cgc gtc gct cac tgc atc gtg gtc gcg tcc	528
Ser Met Met Tyr Pro Glu Arg Val Ala His Cys Ile Val Val Ala Ser	
165 170 175	
aca ccc aag ctg tcg gcg cag aac atc gcg ttc aac gag gtt gcg cgc	576
Thr Pro Lys Leu Ser Ala Gln Asn Ile Ala Phe Asn Glu Val Ala Arg	
180 185 190	
tcg gcg atc ctg tcg gac ccg gac ttc cac ggc ggc aac tac tac gcg	624
Ser Ala Ile Leu Ser Asp Pro Asp Phe His Gly Gly Asn Tyr Tyr Ala	
195 200 205	
cac aac gtt aag ccg aag cgc ggc ctg cgc gtc gcg cgc atg atc ggc	672
His Asn Val Lys Pro Lys Arg Gly Leu Arg Val Ala Arg Met Ile Gly	
210 215 220	
cac atc acg tat ctg tcg gac gac gac atg gcc gag aaa ttc ggc cgc	720
His Ile Thr Tyr Leu Ser Asp Asp Asp Met Ala Glu Lys Phe Gly Arg	
225 230 235 240	

tcg ctg cgg cgc gcg gaa ggc gcg ctg gac gcg tac aac ttc aac ttc 768
 Ser Leu Arg Arg Ala Glu Gly Ala Leu Asp Ala Tyr Asn Phe Asn Phe
 245 250 255
 gac gtg gag ttc gag gtg gag tcg tac ctg cgc tac cag ggc gac aag 816
 Asp Val Glu Phe Glu Val Glu Ser Tyr Leu Arg Tyr Gln Gly Asp Lys
 260 265 270
 ttc gcc gac tac ttc gac gcg aat acg tat ctg ctg atc acc cgc gcg 864
 Phe Ala Asp Tyr Phe Asp Ala Asn Thr Tyr Leu Leu Ile Thr Arg Ala
 275 280 285
 ctc gac tac ttc gat ccg gcc aag gcc ttc gcc ggc gac ctg acg gcc 912
 Leu Asp Tyr Phe Asp Pro Ala Lys Ala Phe Ala Gly Asp Leu Thr Ala
 290 295 300
 gcg gtc gcg cac acc acg gcg aaa tat ctg atc gcc agc ttc acg acc 960
 Ala Val Ala His Thr Thr Ala Lys Tyr Leu Ile Ala Ser Phe Thr Thr
 305 310 315 320
 gac tgg cgc ttc gcg ccg gcc cgc tcg cgt gaa ctg gtg aag gcg ctg 1008
 Asp Trp Arg Phe Ala Pro Ala Arg Ser Arg Glu Leu Val Lys Ala Leu
 325 330 335
 ctc gat cac aag cgc acg gtc acc tac gcg gaa atc gac gcg ccg cac 1056
 Leu Asp His Lys Arg Thr Val Thr Tyr Ala Glu Ile Asp Ala Pro His
 340 345 350
 ggc cac gac gcc ttc ctg ctc gac gac gcg cgc tat cac aac ctg atg 1104
 Gly His Asp Ala Phe Leu Leu Asp Asp Ala Arg Tyr His Asn Leu Met
 355 360 365
 cgc gct tac tac gaa cgt att gcg aac gag gtg aac gca tga 1146
 Arg Ala Tyr Tyr Glu Arg Ile Ala Asn Glu Val Asn Ala
 370 375 380

<210> 20

<211> 381

<212> PRT

<213> Burkholderia cepacia

<400> 20

Met Glu Ser Ile Gly Ile Val Ala Pro Gln Lys Met His Phe Thr Glu
 1 5 10 15
 Pro Leu Pro Leu Gln Asn Gly Ser Ser Leu Ala Gly Tyr Asp Leu Met
 20 25 30
 Val Glu Thr Tyr Gly Thr Leu Asn Ala Ala Arg Ser Asn Ala Val Leu
 35 40 45
 Val Cys His Ala Leu Asn Ala Ser His His Val Ala Gly Val Tyr Ala
 50 55 60
 Asp Asn Pro Arg Asp Ile Gly Trp Trp Asp Asn Met Val Gly Pro Gly
 65 70 75 80
 Lys Pro Leu Asp Thr Asp Lys Phe Phe Val Ile Gly Val Asn Asn Leu
 85 90 95
 Gly Ser Cys Phe Gly Ser Thr Gly Pro Met Ser Ile Asp Pro Ser Thr
 100 105 110

Gly Asn Pro Tyr Gly Ala Thr Phe Pro Val Val Thr Val Glu Asp Trp
115 120 125

Val Asn Ala Gln Ala Arg Val Ala Asp Gln Phe Gly Ile Thr Arg Phe
130 135 140

Ala Ala Val Met Gly Gly Ser Leu Gly Gly Met Gln Ala Leu Ala Trp
145 150 155 160

Ser Met Met Tyr Pro Glu Arg Val Ala His Cys Ile Val Val Ala Ser
165 170 175

Thr Pro Lys Leu Ser Ala Gln Asn Ile Ala Phe Asn Glu Val Ala Arg
180 185 190

Ser Ala Ile Leu Ser Asp Pro Asp Phe His Gly Gly Asn Tyr Tyr Ala
195 200 205

His Asn Val Lys Pro Lys Arg Gly Leu Arg Val Ala Arg Met Ile Gly
210 215 220

His Ile Thr Tyr Leu Ser Asp Asp Asp Met Ala Glu Lys Phe Gly Arg
225 230 235 240

Ser Leu Arg Arg Ala Glu Gly Ala Leu Asp Ala Tyr Asn Phe Asn Phe
245 250 255

Asp Val Glu Phe Glu Val Glu Ser Tyr Leu Arg Tyr Gln Gly Asp Lys
260 265 270

Phe Ala Asp Tyr Phe Asp Ala Asn Thr Tyr Leu Leu Ile Thr Arg Ala
275 280 285

Leu Asp Tyr Phe Asp Pro Ala Lys Ala Phe Ala Gly Asp Leu Thr Ala
290 295 300

Ala Val Ala His Thr Thr Ala Lys Tyr Leu Ile Ala Ser Phe Thr Thr
305 310 315 320

Asp Trp Arg Phe Ala Pro Ala Arg Ser Arg Glu Leu Val Lys Ala Leu
325 330 335

Leu Asp His Lys Arg Thr Val Thr Tyr Ala Glu Ile Asp Ala Pro His
340 345 350

Gly His Asp Ala Phe Leu Leu Asp Asp Ala Arg Tyr His Asn Leu Met
355 360 365

Arg Ala Tyr Tyr Glu Arg Ile Ala Asn Glu Val Asn Ala
370 375 380

<210> 21

<211> 1134

<212> DNA

<213> Nitrosomonas europaea

<220>

<221> CDS

<222> (1) .. (1131)

<223> RNE02005

<400> 21

atg tcc aca caa gat tct gat tcg atc ggc atc gta tcg gca cga cgc	48
Met Ser Thr Gln Asp Ser Asp Ser Ile Gly Ile Val Ser Ala Arg Arg	
1 5 10 15	
gcc cat ttc gac acc ccg ctc agc ctg aaa agc gga gct gta ctg gac	96
Ala His Phe Asp Thr Pro Leu Ser Leu Lys Ser Gly Ala Val Leu Asp	
20 25 30	
agc tac gag ctc gtc tat gaa acc tat ggg gag ctg aat gca gac cga	144
Ser Tyr Glu Leu Val Tyr Glu Thr Tyr Gly Glu Leu Asn Ala Asp Arg	
35 40 45	
tcc aat gca gtg ctg atc tgc cat gct tta tcc ggc aac cac cat gtt	192
Ser Asn Ala Val Leu Ile Cys His Ala Leu Ser Gly Asn His His Val	
50 55 60	
gcc ggt gtt tat gca gat aac ccc aag aat acc gga tgg tgg aac aac	240
Ala Gly Val Tyr Ala Asp Asn Pro Lys Asn Thr Gly Trp Trp Asn Asn	
65 70 75 80	
atg atc ggt ccg ggc aaa ccg gtc gat acc cga aaa ttc ttt gtc atc	288
Met Ile Gly Pro Gly Lys Pro Val Asp Thr Arg Lys Phe Phe Val Ile	
85 90 95	
ggt atc aat aat ctc ggg ggt tgc cat ggc tcc acc ggg ccc atc agc	336
Gly Ile Asn Asn Leu Gly Gly Cys His Gly Ser Thr Gly Pro Ile Ser	
100 105 110	
atc aac gac aag acc ggt aaa cgc ttc ggc ccg gat ttt ccg ctg gta	384
Ile Asn Asp Lys Thr Gly Lys Arg Phe Gly Pro Asp Phe Pro Leu Val	
115 120 125	
acg aca gct gac tgg gca aaa acc tat gtc cgt ttc gcc gat cag ttc	432
Thr Thr Ala Asp Trp Ala Lys Thr Tyr Val Arg Phe Ala Asp Gln Phe	
130 135 140	
agc atc gac tgt ttt gcc gcc gtc atc ggt ggc agt ctg ggc ggg atg	480
Ser Ile Asp Cys Phe Ala Ala Val Ile Gly Gly Ser Leu Gly Gly Met	
145 150 155 160	
tcg gcc atg caa ctg gcg ctc gat gca ccg gaa aga gtt cgt cat gcc	528
Ser Ala Met Gln Leu Ala Leu Asp Ala Pro Glu Arg Val Arg His Ala	
165 170 175	
ata gtg gtt gca gca tcg gcc agg ctg aca gca cag aac atc gct ttc	576
Ile Val Val Ala Ala Ser Ala Arg Leu Thr Ala Gln Asn Ile Ala Phe	
180 185 190	
aat gat gtc gcg cgt cag gcg att ctg acc gac cct gat ttt cac gac	624
Asn Asp Val Ala Arg Gln Ala Ile Leu Thr Asp Pro Asp Phe His Asp	
195 200 205	
ggc gac tat tat tcc cat ggc acc cac ccg cgc aga ggt tta cgc ctt	672
Gly Asp Tyr Tyr Ser His Gly Thr His Pro Arg Arg Gly Leu Arg Leu	
210 215 220	
gcc cgc atg ctt ggc cac atc acc tac ctg tcg gac gac tcc atg gcc	720
Ala Arg Met Leu Gly His Ile Thr Tyr Leu Ser Asp Asp Ser Met Ala	
225 230 235 240	
agc aaa ttc ggc cgt gag tta cgt aac ggc tcg ctt gct ttc aat tat	768
Ser Lys Phe Gly Arg Glu Leu Arg Asn Gly Ser Leu Ala Phe Asn Tyr	

245	250	255	
gat gtg gaa ttc cag atc gaa tcc tat ctg cac cat cag ggc gac aaa			816
Asp Val Glu Phe Gln Ile Glu Ser Tyr Leu His His Gln Gly Asp Lys			
260	265	270	
ttt gcc gac ctg ttc gac gca aac act tat ctg ctg atg acg aag gcg			864
Phe Ala Asp Leu Phe Asp Ala Asn Thr Tyr Leu Leu Met Thr Lys Ala			
275	280	285	
ctc gat tat ttc gat ccg gcc cag gat tac gat ggc aac ctg agt gca			912
Leu Asp Tyr Phe Asp Pro Ala Gln Asp Tyr Asp Gly Asn Leu Ser Ala			
290	295	300	
gcc ttt gcc cgt gca caa gcg gat ttt ctg gta ctt tcc ttt act tcc			960
Ala Phe Ala Arg Ala Gln Ala Asp Phe Leu Val Leu Ser Phe Thr Ser			
305	310	315	320
gac tgg cgt ttt tcc ccg gag cgt tcg cgc gat atc gtc aag gca ctg			1008
Asp Trp Arg Phe Ser Pro Glu Arg Ser Arg Asp Ile Val Lys Ala Leu			
325	330	335	
ctc gac aac aaa ctg aat gtc agt tat gcg gaa att ccc tcc tcg tac			1056
Leu Asp Asn Lys Leu Asn Val Ser Tyr Ala Glu Ile Pro Ser Ser Tyr			
340	345	350	
gga cat gat tcc ttt ctc atg cag gac gac tac tat cac cag ttg ata			1104
Gly His Asp Ser Phe Leu Met Gln Asp Asp Tyr Tyr His Gln Leu Ile			
355	360	365	
cgt gct tac atg aac aat atc gct ctc tag			1134
Arg Ala Tyr Met Asn Asn Ile Ala Leu			
370	375		
<210> 22			
<211> 377			
<212> PRT			
<213> Nitrosomonas europaea			
<400> 22			
Met Ser Thr Gln Asp Ser Asp Ser Ile Gly Ile Val Ser Ala Arg Arg			
1	5	10	15
Ala His Phe Asp Thr Pro Leu Ser Leu Lys Ser Gly Ala Val Leu Asp			
20	25	30	
Ser Tyr Glu Leu Val Tyr Glu Thr Tyr Gly Glu Leu Asn Ala Asp Arg			
35	40	45	
Ser Asn Ala Val Leu Ile Cys His Ala Leu Ser Gly Asn His His Val			
50	55	60	
Ala Gly Val Tyr Ala Asp Asn Pro Lys Asn Thr Gly Trp Trp Asn Asn			
65	70	75	80
Met Ile Gly Pro Gly Lys Pro Val Asp Thr Arg Lys Phe Phe Val Ile			
85	90	95	
Gly Ile Asn Asn Leu Gly Gly Cys His Gly Ser Thr Gly Pro Ile Ser			
100	105	110	
Ile Asn Asp Lys Thr Gly Lys Arg Phe Gly Pro Asp Phe Pro Leu Val			

115					120					125					
Thr	Thr	Ala	Asp	Trp	Ala	Lys	Thr	Tyr	Val	Arg	Phe	Ala	Asp	Gln	Phe
130						135					140				
Ser	Ile	Asp	Cys	Phe	Ala	Ala	Val	Ile	Gly	Gly	Ser	Leu	Gly	Gly	Met
145						150					155				160
Ser	Ala	Met	Gln	Leu	Ala	Leu	Asp	Ala	Pro	Glu	Arg	Val	Arg	His	Ala
				165					170					175	
Ile	Val	Val	Ala	Ala	Ser	Ala	Arg	Leu	Thr	Ala	Gln	Asn	Ile	Ala	Phe
			180					185					190		
Asn	Asp	Val	Ala	Arg	Gln	Ala	Ile	Leu	Thr	Asp	Pro	Asp	Phe	His	Asp
		195					200					205			
Gly	Asp	Tyr	Tyr	Ser	His	Gly	Thr	His	Pro	Arg	Arg	Gly	Leu	Arg	Leu
	210					215					220				
Ala	Arg	Met	Leu	Gly	His	Ile	Thr	Tyr	Leu	Ser	Asp	Asp	Ser	Met	Ala
225						230					235				240
Ser	Lys	Phe	Gly	Arg	Glu	Leu	Arg	Asn	Gly	Ser	Leu	Ala	Phe	Asn	Tyr
			245						250					255	
Asp	Val	Glu	Phe	Gln	Ile	Glu	Ser	Tyr	Leu	His	His	Gln	Gly	Asp	Lys
			260					265					270		
Phe	Ala	Asp	Leu	Phe	Asp	Ala	Asn	Thr	Tyr	Leu	Leu	Met	Thr	Lys	Ala
		275					280					285			
Leu	Asp	Tyr	Phe	Asp	Pro	Ala	Gln	Asp	Tyr	Asp	Gly	Asn	Leu	Ser	Ala
	290						295				300				
Ala	Phe	Ala	Arg	Ala	Gln	Ala	Asp	Phe	Leu	Val	Leu	Ser	Phe	Thr	Ser
305						310					315				320
Asp	Trp	Arg	Phe	Ser	Pro	Glu	Arg	Ser	Arg	Asp	Ile	Val	Lys	Ala	Leu
			325						330					335	
Leu	Asp	Asn	Lys	Leu	Asn	Val	Ser	Tyr	Ala	Glu	Ile	Pro	Ser	Ser	Tyr
		340						345					350		
Gly	His	Asp	Ser	Phe	Leu	Met	Gln	Asp	Asp	Tyr	Tyr	His	Gln	Leu	Ile
	355						360					365			
Arg	Ala	Tyr	Met	Asn	Asn	Ile	Ala	Leu							
	370					375									

<210> 23

<211> 1077

<212> DNA

<213> Haemophilus influenzae

<220>

<221> CDS

<222> (1)..(1074)

<223> RHI02681

<400> 23

atg tct gtg caa aat gta gtg ctt ttt gac aca cag cct tta act ctg 48

Met	Ser	Val	Gln	Asn	Val	Val	Leu	Phe	Asp	Thr	Gln	Pro	Leu	Thr	Leu	
1				5					10					15		
atg	ctt	ggc	ggc	aaa	ctt	tcc	cat	att	aat	gtc	gcg	tat	caa	act	tat	96
Met	Leu	Gly	Gly	Lys	Leu	Ser	His	Ile	Asn	Val	Ala	Tyr	Gln	Thr	Tyr	
		20						25					30			
ggc	acg	ctc	aat	gcc	gaa	aaa	aat	aat	gcg	gta	tta	att	tgc	cac	gct	144
Gly	Thr	Leu	Asn	Ala	Glu	Lys	Asn	Asn	Ala	Val	Leu	Ile	Cys	His	Ala	
		35					40					45				
ttg	act	ggg	gat	gct	gag	cct	tat	ttc	gat	gat	ggg	cga	gat	ggc	tgg	192
Leu	Thr	Gly	Asp	Ala	Glu	Pro	Tyr	Phe	Asp	Asp	Gly	Arg	Asp	Gly	Trp	
	50					55					60					
tgg	cag	aat	ttt	atg	gga	gca	ggg	tta	gca	ttg	gat	acg	gat	cgt	tat	240
Trp	Gln	Asn	Phe	Met	Gly	Ala	Gly	Leu	Ala	Leu	Asp	Thr	Asp	Arg	Tyr	
65					70				75						80	
ttt	ttt	att	agc	tcg	aac	gta	tta	ggg	ggg	tgc	aag	gga	aca	act	ggg	288
Phe	Phe	Ile	Ser	Ser	Asn	Val	Leu	Gly	Gly	Cys	Lys	Gly	Thr	Thr	Gly	
				85					90					95		
cct	tca	tca	att	aat	ccg	caa	acg	ggg	aaa	cct	tat	ggc	agc	caa	ttt	336
Pro	Ser	Ser	Ile	Asn	Pro	Gln	Thr	Gly	Lys	Pro	Tyr	Gly	Ser	Gln	Phe	
			100					105					110			
cct	aat	att	gtt	gtg	caa	gat	att	gtt	aaa	gta	caa	aaa	gcc	ttg	ctt	384
Pro	Asn	Ile	Val	Val	Gln	Asp	Ile	Val	Lys	Val	Gln	Lys	Ala	Leu	Leu	
		115					120					125				
gat	cat	ctt	ggg	att	agc	cat	tta	aaa	gcc	att	att	ggg	gga	tct	ttt	432
Asp	His	Leu	Gly	Ile	Ser	His	Leu	Lys	Ala	Ile	Ile	Gly	Gly	Ser	Phe	
	130					135					140					
ggc	ggc	atg	caa	gcg	aat	caa	tgg	gcg	att	gat	tat	cct	gat	ttt	atg	480
Gly	Gly	Met	Gln	Ala	Asn	Gln	Trp	Ala	Ile	Asp	Tyr	Pro	Asp	Phe	Met	
145					150				155						160	
gat	aat	atc	gtg	aat	ctt	tgc	tca	tcc	att	tat	ttt	agt	gct	gaa	gcc	528
Asp	Asn	Ile	Val	Asn	Leu	Cys	Ser	Ser	Ile	Tyr	Phe	Ser	Ala	Glu	Ala	
			165					170						175		
ata	ggg	ttt	aat	cac	gta	atg	cgt	caa	gcg	gtc	att	aat	gat	ccc	aat	576
Ile	Gly	Phe	Asn	His	Val	Met	Arg	Gln	Ala	Val	Ile	Asn	Asp	Pro	Asn	
			180					185					190			
ttt	aac	ggc	ggc	gat	tat	tat	gag	ggg	aca	ccg	cca	gat	caa	ggg	tta	624
Phe	Asn	Gly	Gly	Asp	Tyr	Tyr	Glu	Gly	Thr	Pro	Pro	Asp	Gln	Gly	Leu	
		195					200					205				
tct	att	gca	cgt	atg	cta	ggg	atg	ctg	act	tac	cgc	acc	gat	tta	caa	672
Ser	Ile	Ala	Arg	Met	Leu	Gly	Met	Leu	Thr	Tyr	Arg	Thr	Asp	Leu	Gln	
	210					215					220					
ctt	gcg	aaa	gcc	ttt	ggg	cgt	gcc	aca	aaa	tca	gat	ggc	agc	ttt	tgg	720
Leu	Ala	Lys	Ala	Phe	Gly	Arg	Ala	Thr	Lys	Ser	Asp	Gly	Ser	Phe	Trp	
225					230				235						240	
ggc	gat	tac	ttt	caa	gtg	gaa	tcc	tat	ctt	tct	tac	caa	ggc	aaa	aaa	768
Gly	Asp	Tyr	Phe	Gln	Val	Glu	Ser	Tyr	Leu	Ser	Tyr	Gln	Gly	Lys	Lys	
				245					250					255		

```

ttc tta gaa cgt ttt gat gcc aat agt tat ttg cat ttg tta cgt gcg      816
Phe Leu Glu Arg Phe Asp Ala Asn Ser Tyr Leu His Leu Leu Arg Ala
      260                      265                      270

ttg gat atg tat gat cca agt ttg ggg tat gac aat gtt aaa gag gca      864
Leu Asp Met Tyr Asp Pro Ser Leu Gly Tyr Asp Asn Val Lys Glu Ala
      275                      280                      285

tta tca cgt att aaa gca cgc tat acg ttg gtt tct gtg aca acg gat      912
Leu Ser Arg Ile Lys Ala Arg Tyr Thr Leu Val Ser Val Thr Thr Asp
      290                      295                      300

caa ctt ttt aaa ccc att gat ctt tat aaa agt aaa cag ctt tta gag      960
Gln Leu Phe Lys Pro Ile Asp Leu Tyr Lys Ser Lys Gln Leu Leu Glu
      305                      310                      315                      320

caa agt gga gtc gat cta cat ttt tat gaa ttc cca tca gat tac gga      1008
Gln Ser Gly Val Asp Leu His Phe Tyr Glu Phe Pro Ser Asp Tyr Gly
      325                      330                      335

cac gat gcg ttt tta gtg gat tat gat cag ttt gaa aaa cga att cga      1056
His Asp Ala Phe Leu Val Asp Tyr Asp Gln Phe Glu Lys Arg Ile Arg
      340                      345                      350

gat ggt ttg gca ggt aat taa      1077
Asp Gly Leu Ala Gly Asn
      355

```

<210> 24

<211> 358

<212> PRT

<213> Haemophilus influenzae

<400> 24

```

Met Ser Val Gln Asn Val Val Leu Phe Asp Thr Gln Pro Leu Thr Leu
  1                      5                      10                      15

```

```

Met Leu Gly Gly Lys Leu Ser His Ile Asn Val Ala Tyr Gln Thr Tyr
      20                      25                      30

```

```

Gly Thr Leu Asn Ala Glu Lys Asn Asn Ala Val Leu Ile Cys His Ala
      35                      40                      45

```

```

Leu Thr Gly Asp Ala Glu Pro Tyr Phe Asp Asp Gly Arg Asp Gly Trp
      50                      55                      60

```

```

Trp Gln Asn Phe Met Gly Ala Gly Leu Ala Leu Asp Thr Asp Arg Tyr
      65                      70                      75                      80

```

```

Phe Phe Ile Ser Ser Asn Val Leu Gly Gly Cys Lys Gly Thr Thr Gly
      85                      90                      95

```

```

Pro Ser Ser Ile Asn Pro Gln Thr Gly Lys Pro Tyr Gly Ser Gln Phe
      100                      105                      110

```

```

Pro Asn Ile Val Val Gln Asp Ile Val Lys Val Gln Lys Ala Leu Leu
      115                      120                      125

```

```

Asp His Leu Gly Ile Ser His Leu Lys Ala Ile Ile Gly Gly Ser Phe
      130                      135                      140

```

```

Gly Gly Met Gln Ala Asn Gln Trp Ala Ile Asp Tyr Pro Asp Phe Met

```

145 150 155 160
 Asp Asn Ile Val Asn Leu Cys Ser Ser Ile Tyr Phe Ser Ala Glu Ala
 165 170 175
 Ile Gly Phe Asn His Val Met Arg Gln Ala Val Ile Asn Asp Pro Asn
 180 185 190
 Phe Asn Gly Gly Asp Tyr Tyr Glu Gly Thr Pro Pro Asp Gln Gly Leu
 195 200 205
 Ser Ile Ala Arg Met Leu Gly Met Leu Thr Tyr Arg Thr Asp Leu Gln
 210 215 220
 Leu Ala Lys Ala Phe Gly Arg Ala Thr Lys Ser Asp Gly Ser Phe Trp
 225 230 235 240
 Gly Asp Tyr Phe Gln Val Glu Ser Tyr Leu Ser Tyr Gln Gly Lys Lys
 245 250 255
 Phe Leu Glu Arg Phe Asp Ala Asn Ser Tyr Leu His Leu Leu Arg Ala
 260 265 270
 Leu Asp Met Tyr Asp Pro Ser Leu Gly Tyr Asp Asn Val Lys Glu Ala
 275 280 285
 Leu Ser Arg Ile Lys Ala Arg Tyr Thr Leu Val Ser Val Thr Thr Asp
 290 295 300
 Gln Leu Phe Lys Pro Ile Asp Leu Tyr Lys Ser Lys Gln Leu Leu Glu
 305 310 315 320
 Gln Ser Gly Val Asp Leu His Phe Tyr Glu Phe Pro Ser Asp Tyr Gly
 325 330 335
 His Asp Ala Phe Leu Val Asp Tyr Asp Gln Phe Glu Lys Arg Ile Arg
 340 345 350
 Asp Gly Leu Ala Gly Asn
 355

<210> 25

<211> 1296

<212> DNA

<213> Halobacterium sp

<220>

<221> CDS

<222> (1)..(1293)

<223> ETX_HALN1

<400> 25

atg ggc cac gat cac gga ctc cac acc aac agt gta cac gcc ggc cag 48
 Met Gly His Asp His Gly Leu His Thr Asn Ser Val His Ala Gly Gln
 1 5 10 15

cgc gtc gac ccg gcc acg ggc gct cgc gcg ccg cca ctc tac cag acc 96
 Arg Val Asp Pro Ala Thr Gly Ala Arg Ala Pro Pro Leu Tyr Gln Thr
 20 25 30

acg tcg tac gcc ttc gag gac agc gcc gat gcc gcc ggc cag ttc gcc 144
 Thr Ser Tyr Ala Phe Glu Asp Ser Ala Asp Ala Ala Gly Gln Phe Ala

35	40	45	
ctt gag cgg gac ggc tac atc tac tcg cgg ctg atg aac ccc acc gtg Leu Glu Arg Asp Gly Tyr Ile Tyr Ser Arg Leu Met Asn Pro Thr Val 50 55 60			192
gag acc ctc cag gac cgc ctc gcc gcc ctc gaa ggc ggc gtc ggc gcg Glu Thr Leu Gln Asp Arg Leu Ala Ala Leu Glu Gly Gly Val Gly Ala 65 70 75 80			240
gtc gcc acc gcg tcc gga atg gcc gcc ctg gac ctc gcg acg ttc ctg Val Ala Thr Ala Ser Gly Met Ala Ala Leu Asp Leu Ala Thr Phe Leu 85 90 95			288
ctg gca cgc gcc ggc gac tcc gtc gtc gcc gcc agc gac ctc tac ggc Leu Ala Arg Ala Gly Asp Ser Val Val Ala Ala Ser Asp Leu Tyr Gly 100 105 110			336
ggc acc gtg acg tac ctc acg cac agc gcc cag cgc cgc ggc gtc gac Gly Thr Val Thr Tyr Leu Thr His Ser Ala Gln Arg Arg Gly Val Asp 115 120 125			384
acg acg ttc gtg gac gtc ctc gac tac gac gcc tac gcc gac gcc atc Thr Thr Phe Val Asp Val Leu Asp Tyr Asp Ala Tyr Ala Asp Ala Ile 130 135 140			432
gac gcc gac acc gcc tac gtg ctc gtc gaa acc gtc ggc aac ccc agc Asp Ala Asp Thr Ala Tyr Val Leu Val Glu Thr Val Gly Asn Pro Ser 145 150 155 160			480
ctg atc acg ccc gac ctc gaa cgc atc gcc gac atc gcc cac gac aac Leu Ile Thr Pro Asp Leu Glu Arg Ile Ala Asp Ile Ala His Asp Asn 165 170 175			528
ggc gtt ccc ctg ctg gtg gac aac acg ttc gcg acc ccc gcg ctg gca Gly Val Pro Leu Leu Val Asp Asn Thr Phe Ala Thr Pro Ala Leu Ala 180 185 190			576
acc ccg atc gac cac ggt gcc gac atc gtc tgg cac tcc acc acc aaa Thr Pro Ile Asp His Gly Ala Asp Ile Val Trp His Ser Thr Thr Lys 195 200 205			624
tgg atc cac ggt gcc ggc acc acc gtc ggc ggc gcg ctc gtc gac gcc Trp Ile His Gly Ala Gly Thr Thr Val Gly Gly Ala Leu Val Asp Ala 210 215 220			672
ggc agc ttc gac tgg gac gcc cac gcc gcc gac tac ccc gag atc gcc Gly Ser Phe Asp Trp Asp Ala His Ala Ala Asp Tyr Pro Glu Ile Ala 225 230 235 240			720
cag gaa aac ccc gcc tac cac ggc gtg acc ttc acc gat cgc ttc ggg Gln Glu Asn Pro Ala Tyr His Gly Val Thr Phe Thr Asp Arg Phe Gly 245 250 255			768
gac gcc gcg ttc acg tac gcc gca atc gcc cgc ggg ctg cgc gat ctg Asp Ala Ala Phe Thr Tyr Ala Ala Ile Ala Arg Gly Leu Arg Asp Leu 260 265 270			816
ggc aac cag cag tcg ccg ttc gac gcc tgg cag acc ctc cag aag ctc Gly Asn Gln Gln Ser Pro Phe Asp Ala Trp Gln Thr Leu Gln Lys Leu 275 280 285			864
gaa acg ctc ccg ctg cgc atg caa caa cac tgc cgg aac gcc cag ctc			912

Glu Thr Leu Pro Leu Arg Met Gln Gln His Cys Arg Asn Ala Gln Leu
 290 295 300
 gtc gcc gaa cac ctc cgg gac cac ccc aac gtg tcg tgg gtc aac tac 960
 Val Ala Glu His Leu Arg Asp His Pro Asn Val Ser Trp Val Asn Tyr
 305 310 315 320
 ccc ggg ctg gcc gac cac gac acc cac gac aac gca acc acc tac ctc 1008
 Pro Gly Leu Ala Asp His Asp Thr His Asp Asn Ala Thr Thr Tyr Leu
 325 330 335
 gat tcg ggc tac gga ggc atg ctc acg ttc ggc gtc gag gac ggc tac 1056
 Asp Ser Gly Tyr Gly Gly Met Leu Thr Phe Gly Val Glu Asp Gly Tyr
 340 345 350
 gag gcc gcc caa tcg gtc acc gag gag acc acg ctt gcc agc ctg ctg 1104
 Glu Ala Ala Gln Ser Val Thr Glu Glu Thr Thr Leu Ala Ser Leu Leu
 355 360 365
 gcg aac gtc ggc gac gcc aaa acg ctc gtg atc cac ccc gcc tcc acc 1152
 Ala Asn Val Gly Asp Ala Lys Thr Leu Val Ile His Pro Ala Ser Thr
 370 375 380
 acc cac cag cag ctc acc ccc gaa gcc cag cgc gcc ggc ggt gtg cgc 1200
 Thr His Gln Gln Leu Thr Pro Glu Ala Gln Arg Ala Gly Gly Val Arg
 385 390 395 400
 ccc gag atg gtg cgg gtg tcg gtc ggc atc gag gac ccc gcc gac atc 1248
 Pro Glu Met Val Arg Val Ser Val Gly Ile Glu Asp Pro Ala Asp Ile
 405 410 415
 gtc gcg gac ctc gaa acc gcc atc gag gcc gcg gtc ggg tcg gcg 1293
 Val Ala Asp Leu Glu Thr Ala Ile Glu Ala Ala Val Gly Ser Ala
 420 425 430
 tag 1296

<210> 26

<211> 431

<212> PRT

<213> Halobacterium sp

<400> 26

Met Gly His Asp His Gly Leu His Thr Asn Ser Val His Ala Gly Gln
 1 5 10 15
 Arg Val Asp Pro Ala Thr Gly Ala Arg Ala Pro Pro Leu Tyr Gln Thr
 20 25 30
 Thr Ser Tyr Ala Phe Glu Asp Ser Ala Asp Ala Ala Gly Gln Phe Ala
 35 40 45
 Leu Glu Arg Asp Gly Tyr Ile Tyr Ser Arg Leu Met Asn Pro Thr Val
 50 55 60
 Glu Thr Leu Gln Asp Arg Leu Ala Ala Leu Glu Gly Gly Val Gly Ala
 65 70 75 80
 Val Ala Thr Ala Ser Gly Met Ala Ala Leu Asp Leu Ala Thr Phe Leu
 85 90 95
 Leu Ala Arg Ala Gly Asp Ser Val Val Ala Ala Ser Asp Leu Tyr Gly

100							105					110				
Gly	Thr	Val	Thr	Tyr	Leu	Thr	His	Ser	Ala	Gln	Arg	Arg	Gly	Val	Asp	
		115					120					125				
Thr	Thr	Phe	Val	Asp	Val	Leu	Asp	Tyr	Asp	Ala	Tyr	Ala	Asp	Ala	Ile	
		130					135					140				
Asp	Ala	Asp	Thr	Ala	Tyr	Val	Leu	Val	Glu	Thr	Val	Gly	Asn	Pro	Ser	
145					150					155			160			
Leu	Ile	Thr	Pro	Asp	Leu	Glu	Arg	Ile	Ala	Asp	Ile	Ala	His	Asp	Asn	
				165					170					175		
Gly	Val	Pro	Leu	Leu	Val	Asp	Asn	Thr	Phe	Ala	Thr	Pro	Ala	Leu	Ala	
				180					185					190		
Thr	Pro	Ile	Asp	His	Gly	Ala	Asp	Ile	Val	Trp	His	Ser	Thr	Thr	Lys	
				195			200					205				
Trp	Ile	His	Gly	Ala	Gly	Thr	Thr	Val	Gly	Gly	Ala	Leu	Val	Asp	Ala	
						215					220					
Gly	Ser	Phe	Asp	Trp	Asp	Ala	His	Ala	Ala	Asp	Tyr	Pro	Glu	Ile	Ala	
225					230					235			240			
Gln	Glu	Asn	Pro	Ala	Tyr	His	Gly	Val	Thr	Phe	Thr	Asp	Arg	Phe	Gly	
				245					250					255		
Asp	Ala	Ala	Phe	Thr	Tyr	Ala	Ala	Ile	Ala	Arg	Gly	Leu	Arg	Asp	Leu	
				260					265					270		
Gly	Asn	Gln	Gln	Ser	Pro	Phe	Asp	Ala	Trp	Gln	Thr	Leu	Gln	Lys	Leu	
						280							285			
Glu	Thr	Leu	Pro	Leu	Arg	Met	Gln	Gln	His	Cys	Arg	Asn	Ala	Gln	Leu	
						295					300					
Val	Ala	Glu	His	Leu	Arg	Asp	His	Pro	Asn	Val	Ser	Trp	Val	Asn	Tyr	
305					310					315			320			
Pro	Gly	Leu	Ala	Asp	His	Asp	Thr	His	Asp	Asn	Ala	Thr	Thr	Tyr	Leu	
				325					330					335		
Asp	Ser	Gly	Tyr	Gly	Gly	Met	Leu	Thr	Phe	Gly	Val	Glu	Asp	Gly	Tyr	
				340					345					350		
Glu	Ala	Ala	Gln	Ser	Val	Thr	Glu	Glu	Thr	Thr	Leu	Ala	Ser	Leu	Leu	
						360							365			
Ala	Asn	Val	Gly	Asp	Ala	Lys	Thr	Leu	Val	Ile	His	Pro	Ala	Ser	Thr	
						375					380					
Thr	His	Gln	Gln	Leu	Thr	Pro	Glu	Ala	Gln	Arg	Ala	Gly	Gly	Val	Arg	
385					390					395			400			
Pro	Glu	Met	Val	Arg	Val	Ser	Val	Gly	Ile	Glu	Asp	Pro	Ala	Asp	Ile	
				405					410					415		
Val	Ala	Asp	Leu	Glu	Thr	Ala	Ile	Glu	Ala	Ala	Val	Gly	Ser	Ala		
				420					425					430		

<210> 27
 <211> 1143
 <212> DNA
 <213> *Thermus thermophilus*

<220>
 <221> CDS
 <222> (1) .. (1140)
 <223> RTT00268

<400> 27
 atg agc gag atc gcc ctc gag gcc tgg ggg gag cac gag gcc ctc ctc 48
 Met Ser Glu Ile Ala Leu Glu Ala Trp Gly Glu His Glu Ala Leu Leu
 1 5 10 15
 ctc aag ccc ccc cgc tcc ccc ctc tcc atc ccc ccg ccc aag ccc cgc 96
 Leu Lys Pro Pro Arg Ser Pro Leu Ser Ile Pro Pro Pro Lys Pro Arg
 20 25 30
 acc gcc gtc ctc ttc ccc agg cgg gag ggg ttc tac acg gag ctc ggg 144
 Thr Ala Val Leu Phe Pro Arg Arg Glu Gly Phe Tyr Thr Glu Leu Gly
 35 40 45
 ggg tac ctc ccc gag gtg cgc ctc cgc ttt gag acc tac ggg acc ctc 192
 Gly Tyr Leu Pro Glu Val Arg Leu Arg Phe Glu Thr Tyr Gly Thr Leu
 50 55 60
 tcc cgc agg cgg gat aac gcc gtc ctc gtc ttc cac gcc ctc acg ggg 240
 Ser Arg Arg Arg Asp Asn Ala Val Leu Val Phe His Ala Leu Thr Gly
 65 70 75 80
 agc gcc cac ctg gcg ggg acc tac gac gag gaa acc ttt aga agc ctc 288
 Ser Ala His Leu Ala Gly Thr Tyr Asp Glu Glu Thr Phe Arg Ser Leu
 85 90 95
 tcc ccc ctg gag cag gcc ttc ggc cgg gaa ggg tgg tgg gac agc ctg 336
 Ser Pro Leu Glu Gln Ala Phe Gly Arg Glu Gly Trp Trp Asp Ser Leu
 100 105 110
 gtg ggg ccc ggg cgg atc ctg gac ccc gcc ctc tac tac gtg gtc tcc 384
 Val Gly Pro Gly Arg Ile Leu Asp Pro Ala Leu Tyr Tyr Val Val Ser
 115 120 125
 gcc aac cac ctg gga agc tgc tac ggc tcc acc ggc ccc ctc tcc cta 432
 Ala Asn His Leu Gly Ser Cys Tyr Gly Ser Thr Gly Pro Leu Ser Leu
 130 135 140
 gac ccc cac acg ggc cgc ccc tac ggg agg gac ttc cct ccc ctt acc 480
 Asp Pro His Thr Gly Arg Pro Tyr Gly Arg Asp Phe Pro Pro Leu Thr
 145 150 155 160
 atc cgc gac ctg gcc cgg gcc cag gcg agg ctt ctg gac cat ctg ggg 528
 Ile Arg Asp Leu Ala Arg Ala Gln Ala Arg Leu Leu Asp His Leu Gly
 165 170 175
 gtg gag aag gcc atc gtc atc ggg ggg agc ctc ggg ggg atg gtg gcc 576
 Val Glu Lys Ala Ile Val Ile Gly Gly Ser Leu Gly Gly Met Val Ala
 180 185 190
 ctg gag ttc gcc ctc atg tac ccg gag agg gtg aag aag ctc gtg gtc 624
 Leu Glu Phe Ala Leu Met Tyr Pro Glu Arg Val Lys Lys Leu Val Val
 195 200 205

ctg gcg gcc ccc gca cgg cac ggc ccc tgg gcc cgg gcc ttc aac cac 672
 Leu Ala Ala Pro Ala Arg His Gly Pro Trp Ala Arg Ala Phe Asn His
 210 215 220
 ctc tcc cgc cag gcc atc ctc caa gac ccc gag tac cag aag ggc aac 720
 Leu Ser Arg Gln Ala Ile Leu Gln Asp Pro Glu Tyr Gln Lys Gly Asn
 225 230 235 240
 cct gcc ccc aag ggc atg gcc ctc gcc cgg gga atc gcc atg atg agc 768
 Pro Ala Pro Lys Gly Met Ala Leu Ala Arg Gly Ile Ala Met Met Ser
 245 250 255
 tac cgg gcc ccc gag ggg ttt gag gcc cgc tgg ggc gcg gag ccc gag 816
 Tyr Arg Ala Pro Glu Gly Phe Glu Ala Arg Trp Gly Ala Glu Pro Glu
 260 265 270
 ctc ggg gaa atc cac ctg gac tac cag ggg gag aag ttc ctc cgg cgc 864
 Leu Gly Glu Ile His Leu Asp Tyr Gln Gly Glu Lys Phe Leu Arg Arg
 275 280 285
 ttc cac gcc gag agc tac ctc gtc ctc tcc cgg gcc atg gac aac cac 912
 Phe His Ala Glu Ser Tyr Leu Val Leu Ser Arg Ala Met Asp Asn His
 290 295 300
 gac gtg ggc cgg ggc cgg ggc ggg gtg gag gag gcc ctg aag cgc ctc 960
 Asp Val Gly Arg Gly Arg Gly Gly Val Glu Glu Ala Leu Lys Arg Leu
 305 310 315 320
 agg gcc atc ccc tcc ctc ttc gtg ggc att gac acc gac ctc ctc tac 1008
 Arg Ala Ile Pro Ser Leu Phe Val Gly Ile Asp Thr Asp Leu Leu Tyr
 325 330 335
 ccc gcc tgg gag gtg agg cag gcg gcc aag gcg gcg ggg gcc cgc tac 1056
 Pro Ala Trp Glu Val Arg Gln Ala Ala Lys Ala Ala Gly Ala Arg Tyr
 340 345 350
 cgg gag atc aaa agc ccc cac ggg cac gac gcc ttc ctc ata gag acc 1104
 Arg Glu Ile Lys Ser Pro His Gly His Asp Ala Phe Leu Ile Glu Thr
 355 360 365
 gac cag gtg gag gag atc ctg gac gcc ttc ctc ccg tag 1143
 Asp Gln Val Glu Glu Ile Leu Asp Ala Phe Leu Pro
 370 375 380

<210> 28

<211> 380

<212> PRT

<213> Thermus thermophilus

<400> 28

Met Ser Glu Ile Ala Leu Glu Ala Trp Gly Glu His Glu Ala Leu Leu
 1 5 10 15

Leu Lys Pro Pro Arg Ser Pro Leu Ser Ile Pro Pro Pro Lys Pro Arg
 20 25 30

Thr Ala Val Leu Phe Pro Arg Arg Glu Gly Phe Tyr Thr Glu Leu Gly
 35 40 45

Gly Tyr Leu Pro Glu Val Arg Leu Arg Phe Glu Thr Tyr Gly Thr Leu
 50 55 60

Ser Arg Arg Arg Asp Asn Ala Val Leu Val Phe His Ala Leu Thr Gly
 65 70 75 80
 Ser Ala His Leu Ala Gly Thr Tyr Asp Glu Glu Thr Phe Arg Ser Leu
 85 90 95
 Ser Pro Leu Glu Gln Ala Phe Gly Arg Glu Gly Trp Trp Asp Ser Leu
 100 105 110
 Val Gly Pro Gly Arg Ile Leu Asp Pro Ala Leu Tyr Tyr Val Val Ser
 115 120 125
 Ala Asn His Leu Gly Ser Cys Tyr Gly Ser Thr Gly Pro Leu Ser Leu
 130 135 140
 Asp Pro His Thr Gly Arg Pro Tyr Gly Arg Asp Phe Pro Pro Leu Thr
 145 150 155 160
 Ile Arg Asp Leu Ala Arg Ala Gln Ala Arg Leu Leu Asp His Leu Gly
 165 170 175
 Val Glu Lys Ala Ile Val Ile Gly Gly Ser Leu Gly Gly Met Val Ala
 180 185 190
 Leu Glu Phe Ala Leu Met Tyr Pro Glu Arg Val Lys Lys Leu Val Val
 195 200 205
 Leu Ala Ala Pro Ala Arg His Gly Pro Trp Ala Arg Ala Phe Asn His
 210 215 220
 Leu Ser Arg Gln Ala Ile Leu Gln Asp Pro Glu Tyr Gln Lys Gly Asn
 225 230 235 240
 Pro Ala Pro Lys Gly Met Ala Leu Ala Arg Gly Ile Ala Met Met Ser
 245 250 255
 Tyr Arg Ala Pro Glu Gly Phe Glu Ala Arg Trp Gly Ala Glu Pro Glu
 260 265 270
 Leu Gly Glu Ile His Leu Asp Tyr Gln Gly Glu Lys Phe Leu Arg Arg
 275 280 285
 Phe His Ala Glu Ser Tyr Leu Val Leu Ser Arg Ala Met Asp Asn His
 290 295 300
 Asp Val Gly Arg Gly Arg Gly Gly Val Glu Glu Ala Leu Lys Arg Leu
 305 310 315 320
 Arg Ala Ile Pro Ser Leu Phe Val Gly Ile Asp Thr Asp Leu Leu Tyr
 325 330 335
 Pro Ala Trp Glu Val Arg Gln Ala Ala Lys Ala Ala Gly Ala Arg Tyr
 340 345 350
 Arg Glu Ile Lys Ser Pro His Gly His Asp Ala Phe Leu Ile Glu Thr
 355 360 365
 Asp Gln Val Glu Glu Ile Leu Asp Ala Phe Leu Pro
 370 375 380

<210> 29

<211> 1005

<212> DNA

<213> *Deinococcus radiodurans*

<220>

<221> CDS

<222> (1) .. (1002)

<223> RDR01287

<400> 29

gtg acc gcc gtg ctc gcg ggc cac gcc tct gcc ctg ctg ctg acc gaa	48
Val Thr Ala Val Leu Ala Gly His Ala Ser Ala Leu Leu Leu Thr Glu	
1 5 10 15	
gaa ccc gac tgt tgc ggg ccg cag acg gtc gtt ctc ttc cgg cgt gag	96
Glu Pro Asp Cys Ser Gly Pro Gln Thr Val Val Leu Phe Arg Arg Glu	
20 25 30	
ccg ctg ctg ctc gac tgc gga cgg gcg ctg agc gac gtg cgg gtg gcc	144
Pro Leu Leu Leu Asp Cys Gly Arg Ala Leu Ser Asp Val Arg Val Ala	
35 40 45	
ttt cac acc tac ggc acg ccg cgc gcc gac gcc acg ctg gtg ctg cac	192
Phe His Thr Tyr Gly Thr Pro Arg Ala Asp Ala Thr Leu Val Leu His	
50 55 60	
gcc ctg acc ggc gac agc gcg gtg cac gag tgg tgg ccc gac ttt ctg	240
Ala Leu Thr Gly Asp Ser Ala Val His Glu Trp Trp Pro Asp Phe Leu	
65 70 75 80	
ggc gcg ggc cgg cca ctg gac ccg gca gac gac tac gtg gtg tgc gcc	288
Gly Ala Gly Arg Pro Leu Asp Pro Ala Asp Asp Tyr Val Val Cys Ala	
85 90 95	
aac gtc ctc ggc ggg tgc gcc ggc acg acg agc gcc gct gaa ctc gcc	336
Asn Val Leu Gly Gly Cys Ala Gly Thr Thr Ser Ala Ala Glu Leu Ala	
100 105 110	
gcc acc tgt tcc gga ccg gtg ccg ctc agc ctg cgc gac atg gcc cgg	384
Ala Thr Cys Ser Gly Pro Val Pro Leu Ser Leu Arg Asp Met Ala Arg	
115 120 125	
gtg ggg cgc gcc ctg ctg gat tct ctc ggc gtg cga cgg gtg cgg gtc	432
Val Gly Arg Ala Leu Leu Asp Ser Leu Gly Val Arg Arg Val Arg Val	
130 135 140	
atc ggc gcg agc atg ggc ggg atg ctc gcc tac gcc tgg ctg ctg gag	480
Ile Gly Ala Ser Met Gly Gly Met Leu Ala Tyr Ala Trp Leu Leu Glu	
145 150 155 160	
tgc ccc gac ctg gtg gaa aag gcc gtg att ata gga gcc ccg gcg cgg	528
Cys Pro Asp Leu Val Glu Lys Ala Val Ile Ile Gly Ala Pro Ala Arg	
165 170 175	
cac tgc ccc tgg gct att gga ctg aac acg gcg gcc cgc agc gcc att	576
His Ser Pro Trp Ala Ile Gly Leu Asn Thr Ala Ala Arg Ser Ala Ile	
180 185 190	
gcc ctc gct ccc ggc ggc gag ggg ctg aag gtg gcg cgc cag att gcc	624
Ala Leu Ala Pro Gly Gly Glu Gly Leu Lys Val Ala Arg Gln Ile Ala	
195 200 205	
atg ctc agt tac cgc agc ccc gaa agc cta agc cgc acg cag gcg ggg	672
Met Leu Ser Tyr Arg Ser Pro Glu Ser Leu Ser Arg Thr Gln Ala Gly	

210

215

220

cag cgc gtg ccg ggg gtg ccc gcc gtt acg tct tac ctg cac tac caa 720
 Gln Arg Val Pro Gly Val Pro Ala Val Thr Ser Tyr Leu His Tyr Gln
 225 230 235 240

ggc gaa aaa ctc gcc gcc cgc ttc gac gag cag acc tac tgc gcc ctc 768
 Gly Glu Lys Leu Ala Ala Arg Phe Asp Glu Gln Thr Tyr Cys Ala Leu
 245 250 255

acc tgg gcg atg gac gcc ttt cag ccg agc agc gcc gac ctc aaa gcg 816
 Thr Trp Ala Met Asp Ala Phe Gln Pro Ser Ser Ala Asp Leu Lys Ala
 260 265 270

gtg cgc gcg ccg gta ctc gtc gtc ggc atc tcc agc gat ctg ctc tac 864
 Val Arg Ala Pro Val Leu Val Val Gly Ile Ser Ser Asp Leu Leu Tyr
 275 280 285

ccc gcc gcc gag gtc cgc gcc tgc gcc gcc gag ctt ccc cac gcc gac 912
 Pro Ala Ala Glu Val Arg Ala Cys Ala Ala Glu Leu Pro His Ala Asp
 290 295 300

tac tgg gaa ctg ggc agc att cac ggc cac gac gcc ttt ttg atg gac 960
 Tyr Trp Glu Leu Gly Ser Ile His Gly His Asp Ala Phe Leu Met Asp
 305 310 315 320

cca cag gac ttg ccg gag cgg gtg ggg gcg ttt ctc agg agt 1002
 Pro Gln Asp Leu Pro Glu Arg Val Gly Ala Phe Leu Arg Ser
 325 330

tga 1005

<210> 30

<211> 334

<212> PRT

<213> *Deinococcus radiodurans*

<400> 30

Val Thr Ala Val Leu Ala Gly His Ala Ser Ala Leu Leu Leu Thr Glu
 1 5 10 15

Glu Pro Asp Cys Ser Gly Pro Gln Thr Val Val Leu Phe Arg Arg Glu
 20 25 30

Pro Leu Leu Leu Asp Cys Gly Arg Ala Leu Ser Asp Val Arg Val Ala
 35 40 45

Phe His Thr Tyr Gly Thr Pro Arg Ala Asp Ala Thr Leu Val Leu His
 50 55 60

Ala Leu Thr Gly Asp Ser Ala Val His Glu Trp Trp Pro Asp Phe Leu
 65 70 75 80

Gly Ala Gly Arg Pro Leu Asp Pro Ala Asp Asp Tyr Val Val Cys Ala
 85 90 95

Asn Val Leu Gly Gly Cys Ala Gly Thr Thr Ser Ala Ala Glu Leu Ala
 100 105 110

Ala Thr Cys Ser Gly Pro Val Pro Leu Ser Leu Arg Asp Met Ala Arg
 115 120 125

Val Gly Arg Ala Leu Leu Asp Ser Leu Gly Val Arg Arg Val Arg Val
 130 135 140

Ile Gly Ala Ser Met Gly Gly Met Leu Ala Tyr Ala Trp Leu Leu Glu
 145 150 155 160

Cys Pro Asp Leu Val Glu Lys Ala Val Ile Ile Gly Ala Pro Ala Arg
 165 170 175

His Ser Pro Trp Ala Ile Gly Leu Asn Thr Ala Ala Arg Ser Ala Ile
 180 185 190

Ala Leu Ala Pro Gly Gly Glu Gly Leu Lys Val Ala Arg Gln Ile Ala
 195 200 205

Met Leu Ser Tyr Arg Ser Pro Glu Ser Leu Ser Arg Thr Gln Ala Gly
 210 215 220

Gln Arg Val Pro Gly Val Pro Ala Val Thr Ser Tyr Leu His Tyr Gln
 225 230 235 240

Gly Glu Lys Leu Ala Ala Arg Phe Asp Glu Gln Thr Tyr Cys Ala Leu
 245 250 255

Thr Trp Ala Met Asp Ala Phe Gln Pro Ser Ser Ala Asp Leu Lys Ala
 260 265 270

Val Arg Ala Pro Val Leu Val Val Gly Ile Ser Ser Asp Leu Leu Tyr
 275 280 285

Pro Ala Ala Glu Val Arg Ala Cys Ala Ala Glu Leu Pro His Ala Asp
 290 295 300

Tyr Trp Glu Leu Gly Ser Ile His Gly His Asp Ala Phe Leu Met Asp
 305 310 315 320

Pro Gln Asp Leu Pro Glu Arg Val Gly Ala Phe Leu Arg Ser
 325 330

<210> 31

<211> 1461

<212> DNA

<213> *Saccharomyces cerevisiae*

<220>

<221> CDS

<222> (1) .. (1458)

<223> RSC08123

<400> 31

atg	tcg	cat	act	tta	aaa	tcg	aaa	acg	ctc	caa	gag	ctg	gac	att	gag	48
Met	Ser	His	Thr	Leu	Lys	Ser	Lys	Thr	Leu	Gln	Glu	Leu	Asp	Ile	Glu	
1				5					10					15		
gag	att	aag	gaa	act	aac	cca	ttg	ctc	aaa	cta	gtt	caa	ggg	cag	agg	96
Glu	Ile	Lys	Glu	Thr	Asn	Pro	Leu	Leu	Lys	Leu	Val	Gln	Gly	Gln	Arg	
			20					25					30			
att	gtt	caa	gtt	ccg	gaa	cta	gtg	ctt	gag	tct	ggc	gtg	gtc	ata	aat	144
Ile	Val	Gln	Val	Pro	Glu	Leu	Val	Leu	Glu	Ser	Gly	Val	Val	Ile	Asn	
			35				40					45				

aat ttc cct att gct tat aag acg tgg ggt aca ctg aat gaa gct ggt	192
Asn Phe Pro Ile Ala Tyr Lys Thr Trp Gly Thr Leu Asn Glu Ala Gly	
50 55 60	
gat aat gtt ctg gta att tgt cat gcc ttg act ggg tcc gca gat gtt	240
Asp Asn Val Leu Val Ile Cys His Ala Leu Thr Gly Ser Ala Asp Val	
65 70 75 80	
gct gac tgg tgg ggc cct ctt ctg ggt aac gac tta gca ttc gac cca	288
Ala Asp Trp Trp Gly Pro Leu Leu Gly Asn Asp Leu Ala Phe Asp Pro	
85 90 95	
tca agg ttt ttt atc ata tgt tta aac tct atg ggc tct cca tat ggg	336
Ser Arg Phe Phe Ile Ile Cys Leu Asn Ser Met Gly Ser Pro Tyr Gly	
100 105 110	
tct ttt tcg cca tta acg ata aat gag gag acg ggc gtt aga tat gga	384
Ser Phe Ser Pro Leu Thr Ile Asn Glu Glu Thr Gly Val Arg Tyr Gly	
115 120 125	
ccc gaa ttc cca tta tgt act gtg cgc gat gac gtt aga gct cac aga	432
Pro Glu Phe Pro Leu Cys Thr Val Arg Asp Asp Val Arg Ala His Arg	
130 135 140	
att gtt ctg gat tct ctg gga gta aag tca ata gcc tgt gtt att ggt	480
Ile Val Leu Asp Ser Leu Gly Val Lys Ser Ile Ala Cys Val Ile Gly	
145 150 155 160	
ggc tct atg ggg ggg atg ctg agt ttg gaa tgg gct gcc atg tat ggt	528
Gly Ser Met Gly Gly Met Leu Ser Leu Glu Trp Ala Ala Met Tyr Gly	
165 170 175	
aag gaa tat gtg aag aat atg gtt gct ctg gcg aca tca gca aga cat	576
Lys Glu Tyr Val Lys Asn Met Val Ala Leu Ala Thr Ser Ala Arg His	
180 185 190	
tct gcc tgg tgc ata tcg tgg tct gag gct caa aga caa tcg att tac	624
Ser Ala Trp Cys Ile Ser Trp Ser Glu Ala Gln Arg Gln Ser Ile Tyr	
195 200 205	
tca gat ccc aac tac ttg gac ggg tac tat ccg gta gag gag caa cct	672
Ser Asp Pro Asn Tyr Leu Asp Gly Tyr Tyr Pro Val Glu Glu Gln Pro	
210 215 220	
gtg gcc gga cta tcg gct gca cgt atg tct gca ttg ttg acg tac agg	720
Val Ala Gly Leu Ser Ala Ala Arg Met Ser Ala Leu Leu Thr Tyr Arg	
225 230 235 240	
aca aga aac agt ttc gag aac aaa ttc tcc aga aga tct cct tca ata	768
Thr Arg Asn Ser Phe Glu Asn Lys Phe Ser Arg Arg Ser Pro Ser Ile	
245 250 255	
gca caa caa caa aaa gct caa agg gag gag aca cgc aaa cca tct act	816
Ala Gln Gln Gln Lys Ala Gln Arg Glu Glu Thr Arg Lys Pro Ser Thr	
260 265 270	
gtc agc gaa cac tcc cta caa atc cac aat gat ggg tat aaa aca aaa	864
Val Ser Glu His Ser Leu Gln Ile His Asn Asp Gly Tyr Lys Thr Lys	
275 280 285	
gcc agc act gcc atc gct ggc att tct ggg caa aaa ggt caa agc gtg	912
Ala Ser Thr Ala Ile Ala Gly Ile Ser Gly Gln Lys Gly Gln Ser Val	
290 295 300	

gtg tcc acc gca tct tct tcg gat tca ttg aat tct tca aca tcg atg 960
 Val Ser Thr Ala Ser Ser Ser Asp Ser Leu Asn Ser Ser Thr Ser Met
 305 310 315 320
 act tcg gta agt tct gta acg ggt gaa gtg aag gac ata aag cct gcg 1008
 Thr Ser Val Ser Ser Val Thr Gly Glu Val Lys Asp Ile Lys Pro Ala
 325 330 335
 cag acg tat ttt tct gca caa agt tac ttg agg tac cag ggc aca aag 1056
 Gln Thr Tyr Phe Ser Ala Gln Ser Tyr Leu Arg Tyr Gln Gly Thr Lys
 340 345 350
 ttc atc aat agg ttc gac gcc aat tgt tac att gcc atc aca cgt aaa 1104
 Phe Ile Asn Arg Phe Asp Ala Asn Cys Tyr Ile Ala Ile Thr Arg Lys
 355 360 365
 ctg gat acg cac gat ttg gca aga gac aga gta gat gac atc act gag 1152
 Leu Asp Thr His Asp Leu Ala Arg Asp Arg Val Asp Asp Ile Thr Glu
 370 375 380
 gtc ctt tct acc atc caa caa cca tcc ctg atc atc ggt atc caa tct 1200
 Val Leu Ser Thr Ile Gln Gln Pro Ser Leu Ile Ile Gly Ile Gln Ser
 385 390 395 400
 gat gga ctg ttc aca tat tca gaa caa gaa ttt ttg gct gag cac ata 1248
 Asp Gly Leu Phe Thr Tyr Ser Glu Gln Glu Phe Leu Ala Glu His Ile
 405 410 415
 ccg aag tcg caa tta gaa aaa att gaa tct ccc gaa ggc cac gat gcc 1296
 Pro Lys Ser Gln Leu Glu Lys Ile Glu Ser Pro Glu Gly His Asp Ala
 420 425 430
 ttc cta ttg gag ttt aag ctg ata aac aaa ctg ata gta caa ttt tta 1344
 Phe Leu Leu Glu Phe Lys Leu Ile Asn Lys Leu Ile Val Gln Phe Leu
 435 440 445
 aaa acc aac tgc aag gcc att acc gat gcc gct cca aga gct tgg gga 1392
 Lys Thr Asn Cys Lys Ala Ile Thr Asp Ala Ala Pro Arg Ala Trp Gly
 450 455 460
 ggt gac gtt ggt aac gat gaa acg aag acg tct gtc ttt ggt gag gcc 1440
 Gly Asp Val Gly Asn Asp Glu Thr Lys Thr Ser Val Phe Gly Glu Ala
 465 470 475 480
 gaa gaa gtt acc aac tgg tag 1461
 Glu Glu Val Thr Asn Trp
 485

<210> 32

<211> 486

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 32

Met Ser His Thr Leu Lys Ser Lys Thr Leu Gln Glu Leu Asp Ile Glu
 1 5 10 15

Glu Ile Lys Glu Thr Asn Pro Leu Leu Lys Leu Val Gln Gly Gln Arg
 20 25 30

Ile Val Gln Val Pro Glu Leu Val Leu Glu Ser Gly Val Val Ile Asn

35					40					45					
Asn	Phe	Pro	Ile	Ala	Tyr	Lys	Thr	Trp	Gly	Thr	Leu	Asn	Glu	Ala	Gly
50						55					60				
Asp	Asn	Val	Leu	Val	Ile	Cys	His	Ala	Leu	Thr	Gly	Ser	Ala	Asp	Val
65					70					75					80
Ala	Asp	Trp	Trp	Gly	Pro	Leu	Leu	Gly	Asn	Asp	Leu	Ala	Phe	Asp	Pro
				85					90					95	
Ser	Arg	Phe	Phe	Ile	Ile	Cys	Leu	Asn	Ser	Met	Gly	Ser	Pro	Tyr	Gly
		100						105					110		
Ser	Phe	Ser	Pro	Leu	Thr	Ile	Asn	Glu	Glu	Thr	Gly	Val	Arg	Tyr	Gly
		115					120					125			
Pro	Glu	Phe	Pro	Leu	Cys	Thr	Val	Arg	Asp	Asp	Val	Arg	Ala	His	Arg
130						135					140				
Ile	Val	Leu	Asp	Ser	Leu	Gly	Val	Lys	Ser	Ile	Ala	Cys	Val	Ile	Gly
145					150					155					160
Gly	Ser	Met	Gly	Gly	Met	Leu	Ser	Leu	Glu	Trp	Ala	Ala	Met	Tyr	Gly
			165						170					175	
Lys	Glu	Tyr	Val	Lys	Asn	Met	Val	Ala	Leu	Ala	Thr	Ser	Ala	Arg	His
			180					185					190		
Ser	Ala	Trp	Cys	Ile	Ser	Trp	Ser	Glu	Ala	Gln	Arg	Gln	Ser	Ile	Tyr
		195					200					205			
Ser	Asp	Pro	Asn	Tyr	Leu	Asp	Gly	Tyr	Tyr	Pro	Val	Glu	Glu	Gln	Pro
210						215					220				
Val	Ala	Gly	Leu	Ser	Ala	Ala	Arg	Met	Ser	Ala	Leu	Leu	Thr	Tyr	Arg
225					230					235					240
Thr	Arg	Asn	Ser	Phe	Glu	Asn	Lys	Phe	Ser	Arg	Arg	Ser	Pro	Ser	Ile
				245					250					255	
Ala	Gln	Gln	Gln	Lys	Ala	Gln	Arg	Glu	Glu	Thr	Arg	Lys	Pro	Ser	Thr
			260					265					270		
Val	Ser	Glu	His	Ser	Leu	Gln	Ile	His	Asn	Asp	Gly	Tyr	Lys	Thr	Lys
		275					280					285			
Ala	Ser	Thr	Ala	Ile	Ala	Gly	Ile	Ser	Gly	Gln	Lys	Gly	Gln	Ser	Val
290						295					300				
Val	Ser	Thr	Ala	Ser	Ser	Ser	Asp	Ser	Leu	Asn	Ser	Ser	Thr	Ser	Met
305					310					315					320
Thr	Ser	Val	Ser	Ser	Val	Thr	Gly	Glu	Val	Lys	Asp	Ile	Lys	Pro	Ala
				325					330					335	
Gln	Thr	Tyr	Phe	Ser	Ala	Gln	Ser	Tyr	Leu	Arg	Tyr	Gln	Gly	Thr	Lys
			340					345					350		
Phe	Ile	Asn	Arg	Phe	Asp	Ala	Asn	Cys	Tyr	Ile	Ala	Ile	Thr	Arg	Lys
		355					360					365			
Leu	Asp	Thr	His	Asp	Leu	Ala	Arg	Asp	Arg	Val	Asp	Asp	Ile	Thr	Glu

370

375

380

Val Leu Ser Thr Ile Gln Gln Pro Ser Leu Ile Ile Gly Ile Gln Ser
 385 390 395 400

Asp Gly Leu Phe Thr Tyr Ser Glu Gln Glu Phe Leu Ala Glu His Ile
 405 410 415

Pro Lys Ser Gln Leu Glu Lys Ile Glu Ser Pro Glu Gly His Asp Ala
 420 425 430

Phe Leu Leu Glu Phe Lys Leu Ile Asn Lys Leu Ile Val Gln Phe Leu
 435 440 445

Lys Thr Asn Cys Lys Ala Ile Thr Asp Ala Ala Pro Arg Ala Trp Gly
 450 455 460

Gly Asp Val Gly Asn Asp Glu Thr Lys Thr Ser Val Phe Gly Glu Ala
 465 470 475 480

Glu Glu Val Thr Asn Trp
 485

<210> 33

<211> 1470

<212> DNA

<213> Schizosaccharomyces pombe

<220>

<221> CDS

<222> (1) .. (1467)

<223> RSO01936

<400> 33

atg gaa tct caa tct ccg att gaa tca att gtc ttt act gac tcc tgt 48
 Met Glu Ser Gln Ser Pro Ile Glu Ser Ile Val Phe Thr Asp Ser Cys
 1 5 10 15

cat ccg tct cag caa gaa aat aaa ttt gtt cag ctt att tca gat caa 96
 His Pro Ser Gln Gln Glu Asn Lys Phe Val Gln Leu Ile Ser Asp Gln
 20 25 30

aaa att gca att gtt ccc aaa ttt acg ttg gag tgt ggc gac atc ctt 144
 Lys Ile Ala Ile Val Pro Lys Phe Thr Leu Glu Cys Gly Asp Ile Leu
 35 40 45

tac gat gtt ccc gtt gcc ttc aag act tgg ggt act ttg aat aaa gaa 192
 Tyr Asp Val Pro Val Ala Phe Lys Thr Trp Gly Thr Leu Asn Lys Glu
 50 55 60

gga aac aat tgt ctt ctt ctt tgt cat gct tta agt ggt tct gct gat 240
 Gly Asn Asn Cys Leu Leu Leu Cys His Ala Leu Ser Gly Ser Ala Asp
 65 70 75 80

gct gga gat tgg tgg ggt cct tta ctc ggt cct ggt cgt gcg ttt gat 288
 Ala Gly Asp Trp Trp Gly Pro Leu Leu Gly Pro Gly Arg Ala Phe Asp
 85 90 95

cca tca cat ttc ttt atc gta tgc ctt aat tct ctt ggt agc cca tac 336
 Pro Ser His Phe Phe Ile Val Cys Leu Asn Ser Leu Gly Ser Pro Tyr
 100 105 110

gga agc gcc tct cct gtt aca tgg aac gct gag act cat agt gtt tat	384
Gly Ser Ala Ser Pro Val Thr Trp Asn Ala Glu Thr His Ser Val Tyr	
115 120 125	
ggg cca gaa ttt cct tta gca acc ata cgt gat gat gta aac atc cat	432
Gly Pro Glu Phe Pro Leu Ala Thr Ile Arg Asp Asp Val Asn Ile His	
130 135 140	
aaa ctt att tta caa aga ttg ggt gta aag caa att gct atg gca gta	480
Lys Leu Ile Leu Gln Arg Leu Gly Val Lys Gln Ile Ala Met Ala Val	
145 150 155 160	
ggt ggc tcc atg ggt ggt atg ctg gtt ttg gag tgg gca ttt gat aag	528
Gly Gly Ser Met Gly Gly Met Leu Val Leu Glu Trp Ala Phe Asp Lys	
165 170 175	
gaa ttt gtg cga tca att gtt ccc att tct acc tct ctt cgt cat tcc	576
Glu Phe Val Arg Ser Ile Val Pro Ile Ser Thr Ser Leu Arg His Ser	
180 185 190	
gcg tgg tgc att agc tgg tct gaa gcg caa cgc cag agt ata tat tct	624
Ala Trp Cys Ile Ser Trp Ser Glu Ala Gln Arg Gln Ser Ile Tyr Ser	
195 200 205	
gac cct aag ttt aat gat gga tac tac ggc ata gac gat cag cct gta	672
Asp Pro Lys Phe Asn Asp Gly Tyr Tyr Gly Ile Asp Asp Gln Pro Val	
210 215 220	
agt ggc ctt gga gct gct cgt atg tct gcc ttg ttg aca tat cgc tcc	720
Ser Gly Leu Gly Ala Ala Arg Met Ser Ala Leu Leu Thr Tyr Arg Ser	
225 230 235 240	
aaa tgt tct ttc gaa cgt cgc ttt gcc cgt act gtt cct gat gcg tct	768
Lys Cys Ser Phe Glu Arg Arg Phe Ala Arg Thr Val Pro Asp Ala Ser	
245 250 255	
cgt cac ccc tat cca gat cgt tta ccc act cct ctc acg ccc agt aat	816
Arg His Pro Tyr Pro Asp Arg Leu Pro Thr Pro Leu Thr Pro Ser Asn	
260 265 270	
gca cat tgg gtc gtt cac aac gaa gga aac cgt aat cgc cgt gaa cga	864
Ala His Trp Val Val His Asn Glu Gly Asn Arg Asn Arg Arg Glu Arg	
275 280 285	
cct tgt cga tcc aat gga tca tca cct act tct gaa agt gct tta aat	912
Pro Cys Arg Ser Asn Gly Ser Ser Pro Thr Ser Glu Ser Ala Leu Asn	
290 295 300	
tcc ccc gcc tct tct gtc tcg tct tta cct tct tta ggt gcc tct cag	960
Ser Pro Ala Ser Ser Val Ser Ser Leu Pro Ser Leu Gly Ala Ser Gln	
305 310 315 320	
act aca gac agt tct tcc ctt aac cag agt tcg tta tta aga cgt cct	1008
Thr Thr Asp Ser Ser Ser Leu Asn Gln Ser Ser Leu Leu Arg Arg Pro	
325 330 335	
gct aat act tac ttc tct gcg caa tcg tat tta cgt tac caa gcg aag	1056
Ala Asn Thr Tyr Phe Ser Ala Gln Ser Tyr Leu Arg Tyr Gln Ala Lys	
340 345 350	
aag ttt gta agt cgc ttt gat gct aat tgt tac att tcg att act aaa	1104
Lys Phe Val Ser Arg Phe Asp Ala Asn Cys Tyr Ile Ser Ile Thr Lys	
355 360 365	

aag ttg gac acc cat gat att act cgt gga cgc ggt tca gac tct cct 1152
Lys Leu Asp Thr His Asp Ile Thr Arg Gly Arg Gly Ser Asp Ser Pro
370 375 380

aag gaa gtc atg aag gat ttg tct tta ccc gta ctc gta ctc ggt att 1200
Lys Glu Val Met Lys Asp Leu Ser Leu Pro Val Leu Val Leu Gly Ile
385 390 395 400

gaa agc gat ggt ctt ttc aca ttt gac gaa caa gtt gaa att gcc aaa 1248
Glu Ser Asp Gly Leu Phe Thr Phe Asp Glu Gln Val Glu Ile Ala Lys
405 410 415

tct ttt ccc aat gct acc ttg gaa aaa att att tcg gcc gaa ggc cac 1296
Ser Phe Pro Asn Ala Thr Leu Glu Lys Ile Ile Ser Ala Glu Gly His
420 425 430

gac ggt ttt ttg ctt gag ttt act caa gta aac tca cat att caa aaa 1344
Asp Gly Phe Leu Leu Glu Phe Thr Gln Val Asn Ser His Ile Gln Lys
435 440 445

ttc caa aag gaa cat tta att gat atc atg tct caa act aat tcc ttt 1392
Phe Gln Lys Glu His Leu Ile Asp Ile Met Ser Gln Thr Asn Ser Phe
450 455 460

gag cga ctt gat tcc caa gtt aat gat acc aac cgc gaa agc gtt ttt 1440
Glu Arg Leu Asp Ser Gln Val Asn Asp Thr Asn Arg Glu Ser Val Phe
465 470 475 480

gga gaa atg gaa gac ata acc tcc tgg taa 1470
Gly Glu Met Glu Asp Ile Thr Ser Trp
485

<210> 34

<211> 489

<212> PRT

<213> Schizosaccharomyces pombe

<400> 34

Met Glu Ser Gln Ser Pro Ile Glu Ser Ile Val Phe Thr Asp Ser Cys
1 5 10 15

His Pro Ser Gln Gln Glu Asn Lys Phe Val Gln Leu Ile Ser Asp Gln
20 25 30

Lys Ile Ala Ile Val Pro Lys Phe Thr Leu Glu Cys Gly Asp Ile Leu
35 40 45

Tyr Asp Val Pro Val Ala Phe Lys Thr Trp Gly Thr Leu Asn Lys Glu
50 55 60

Gly Asn Asn Cys Leu Leu Leu Cys His Ala Leu Ser Gly Ser Ala Asp
65 70 75 80

Ala Gly Asp Trp Trp Gly Pro Leu Leu Gly Pro Gly Arg Ala Phe Asp
85 90 95

Pro Ser His Phe Phe Ile Val Cys Leu Asn Ser Leu Gly Ser Pro Tyr
100 105 110

Gly Ser Ala Ser Pro Val Thr Trp Asn Ala Glu Thr His Ser Val Tyr
115 120 125

Gly Pro Glu Phe Pro Leu Ala Thr Ile Arg Asp Asp Val Asn Ile His
 130 135 140
 Lys Leu Ile Leu Gln Arg Leu Gly Val Lys Gln Ile Ala Met Ala Val
 145 150 155 160
 Gly Gly Ser Met Gly Gly Met Leu Val Leu Glu Trp Ala Phe Asp Lys
 165 170 175
 Glu Phe Val Arg Ser Ile Val Pro Ile Ser Thr Ser Leu Arg His Ser
 180 185 190
 Ala Trp Cys Ile Ser Trp Ser Glu Ala Gln Arg Gln Ser Ile Tyr Ser
 195 200 205
 Asp Pro Lys Phe Asn Asp Gly Tyr Tyr Gly Ile Asp Asp Gln Pro Val
 210 215 220
 Ser Gly Leu Gly Ala Ala Arg Met Ser Ala Leu Leu Thr Tyr Arg Ser
 225 230 235 240
 Lys Cys Ser Phe Glu Arg Arg Phe Ala Arg Thr Val Pro Asp Ala Ser
 245 250 255
 Arg His Pro Tyr Pro Asp Arg Leu Pro Thr Pro Leu Thr Pro Ser Asn
 260 265 270
 Ala His Trp Val Val His Asn Glu Gly Asn Arg Asn Arg Arg Glu Arg
 275 280 285
 Pro Cys Arg Ser Asn Gly Ser Ser Pro Thr Ser Glu Ser Ala Leu Asn
 290 295 300
 Ser Pro Ala Ser Ser Val Ser Ser Leu Pro Ser Leu Gly Ala Ser Gln
 305 310 315 320
 Thr Thr Asp Ser Ser Ser Leu Asn Gln Ser Ser Leu Leu Arg Arg Pro
 325 330 335
 Ala Asn Thr Tyr Phe Ser Ala Gln Ser Tyr Leu Arg Tyr Gln Ala Lys
 340 345 350
 Lys Phe Val Ser Arg Phe Asp Ala Asn Cys Tyr Ile Ser Ile Thr Lys
 355 360 365
 Lys Leu Asp Thr His Asp Ile Thr Arg Gly Arg Gly Ser Asp Ser Pro
 370 375 380
 Lys Glu Val Met Lys Asp Leu Ser Leu Pro Val Leu Val Leu Gly Ile
 385 390 395 400
 Glu Ser Asp Gly Leu Phe Thr Phe Asp Glu Gln Val Glu Ile Ala Lys
 405 410 415
 Ser Phe Pro Asn Ala Thr Leu Glu Lys Ile Ile Ser Ala Glu Gly His
 420 425 430
 Asp Gly Phe Leu Leu Glu Phe Thr Gln Val Asn Ser His Ile Gln Lys
 435 440 445
 Phe Gln Lys Glu His Leu Ile Asp Ile Met Ser Gln Thr Asn Ser Phe
 450 455 460

Glu Arg Leu Asp Ser Gln Val Asn Asp Thr Asn Arg Glu Ser Val Phe
 465 470 475 480

Gly Glu Met Glu Asp Ile Thr Ser Trp
 485

<210> 35

<211> 1113

<212> DNA

<213> Xylella almond

<220>

<221> CDS

<222> (1)..(1110)

<223> RXFX01562

<400> 35

atg acc gaa ttt atc cct ccg ggc agc cta ttc cat gcg ctc tcc tct 48
 Met Thr Glu Phe Ile Pro Pro Gly Ser Leu Phe His Ala Leu Ser Ser
 1 5 10 15

cca ttt gcg atg aag cgt ggc gga caa ctc cac cac gcc cgc atc gct 96
 Pro Phe Ala Met Lys Arg Gly Gly Gln Leu His His Ala Arg Ile Ala
 20 25 30

tac gaa aca tgg ggc cgc ctc aat gcc agc gcc acc aat gcc att ctg 144
 Tyr Glu Thr Trp Gly Arg Leu Asn Ala Ser Ala Thr Asn Ala Ile Leu
 35 40 45

atc atg cct ggc tta tca ccc aat gca cat gcc gca cac cat gac agc 192
 Ile Met Pro Gly Leu Ser Pro Asn Ala His Ala Ala His His Asp Ser
 50 55 60

aat gct gag cca ggc tgg tgg gag tca atg cta ggt cca ggc aaa ccc 240
 Asn Ala Glu Pro Gly Trp Trp Glu Ser Met Leu Gly Pro Gly Lys Pro
 65 70 75 80

atc gac aca gac cgt tgg ttc gtg atc tgt gtc aac tca ctt ggt agc 288
 Ile Asp Thr Asp Arg Trp Phe Val Ile Cys Val Asn Ser Leu Gly Ser
 85 90 95

tgc aaa gga tcg act ggc cct gca tcg tac aac ccc atc acg cag gcc 336
 Cys Lys Gly Ser Thr Gly Pro Ala Ser Tyr Asn Pro Ile Thr Gln Ala
 100 105 110

atg tat cgt ttg gac ttt cca gca ctg tca atc gaa gac ggg gcc aac 384
 Met Tyr Arg Leu Asp Phe Pro Ala Leu Ser Ile Glu Asp Gly Ala Asn
 115 120 125

tcc gca att gaa gtg gta cat gca ctg ggc atc aag caa ctt gcc agc 432
 Ser Ala Ile Glu Val Val His Ala Leu Gly Ile Lys Gln Leu Ala Ser
 130 135 140

ctg atc ggc aat tca atg ggc ggc atg acg gca ctg gcc atc ctg ctg 480
 Leu Ile Gly Asn Ser Met Gly Gly Met Thr Ala Leu Ala Ile Leu Leu
 145 150 155 160

tta cat cca gat ata gcc cgc agc cac atc aac atc tca ggc agc gcg 528
 Leu His Pro Asp Ile Ala Arg Ser His Ile Asn Ile Ser Gly Ser Ala
 165 170 175

cag gca tta ccg ttt tcc atc gcc att cgc tcg cta caa cgc gag gcg 576
 Gln Ala Leu Pro Phe Ser Ile Ala Ile Arg Ser Leu Gln Arg Glu Ala
 180 185 190

atc cgc ctg gac ccc cat tgg agg cag gga gac tac gac gac acc cac 624
 Ile Arg Leu Asp Pro His Trp Arg Gln Gly Asp Tyr Asp Asp Thr His
 195 200 205

tac ccg gaa tcg ggg cta cgc atc gca cgc aaa ctt ggg gtg atc acc 672
 Tyr Pro Glu Ser Gly Leu Arg Ile Ala Arg Lys Leu Gly Val Ile Thr
 210 215 220

tac cgc tcc gcg ctg gaa tgg gac ggg cgt ttt ggc cgg gta cgc ttg 720
 Tyr Arg Ser Ala Leu Glu Trp Asp Gly Arg Phe Gly Arg Val Arg Leu
 225 230 235 240

gat tcg gac caa acc aac gac aca cca ttc gga ctg gaa ttc caa att 768
 Asp Ser Asp Gln Thr Asn Asp Thr Pro Phe Gly Leu Glu Phe Gln Ile
 245 250 255

gaa aac tac ttg gaa agc cat gca cac cgc ttc gtg cac acc ttc gac 816
 Glu Asn Tyr Leu Glu Ser His Ala His Arg Phe Val His Thr Phe Asp
 260 265 270

cca aac tgc tac ctg tac ctg agc cgc tcc atg gac tgg ttc gac gtg 864
 Pro Asn Cys Tyr Leu Tyr Leu Ser Arg Ser Met Asp Trp Phe Asp Val
 275 280 285

gcc gag tac gcc aat gga gac att ctt gcc ggg ctg gcc agg atc cga 912
 Ala Glu Tyr Ala Asn Gly Asp Ile Leu Ala Gly Leu Ala Arg Ile Arg
 290 295 300

atc caa cgc gca ctc gcc atc ggt agc cat acc gac atc ctc ttt cca 960
 Ile Gln Arg Ala Leu Ala Ile Gly Ser His Thr Asp Ile Leu Phe Pro
 305 310 315 320

ata caa cag caa caa caa att gcc gaa ggg cta cgc cgt ggc ggt aca 1008
 Ile Gln Gln Gln Gln Gln Ile Ala Glu Gly Leu Arg Arg Gly Gly Thr
 325 330 335

cac gcc acc ttc ctg ggc ctt gac tca ccg cag ggg cat gat gcg ttc 1056
 His Ala Thr Phe Leu Gly Leu Asp Ser Pro Gln Gly His Asp Ala Phe
 340 345 350

ctt gtg gat atc gca aga ttt ggc cct cca gtg aag gaa ttt ctg gac 1104
 Leu Val Asp Ile Ala Arg Phe Gly Pro Pro Val Lys Glu Phe Leu Asp
 355 360 365

gaa ctg tga 1113
 Glu Leu
 370

<210> 36

<211> 370

<212> PRT

<213> Xylella almond

<400> 36

Met Thr Glu Phe Ile Pro Pro Gly Ser Leu Phe His Ala Leu Ser Ser
 1 5 10 15

Pro Phe Ala Met Lys Arg Gly Gly Gln Leu His His Ala Arg Ile Ala

20					25					30						
Tyr	Glu	Thr	Trp	Gly	Arg	Leu	Asn	Ala	Ser	Ala	Thr	Asn	Ala	Ile	Leu	
35					40					45						
Ile	Met	Pro	Gly	Leu	Ser	Pro	Asn	Ala	His	Ala	Ala	His	His	Asp	Ser	
50					55					60						
Asn	Ala	Glu	Pro	Gly	Trp	Trp	Glu	Ser	Met	Leu	Gly	Pro	Gly	Lys	Pro	
65					70					75					80	
Ile	Asp	Thr	Asp	Arg	Trp	Phe	Val	Ile	Cys	Val	Asn	Ser	Leu	Gly	Ser	
85					90					95						
Cys	Lys	Gly	Ser	Thr	Gly	Pro	Ala	Ser	Tyr	Asn	Pro	Ile	Thr	Gln	Ala	
100					105					110						
Met	Tyr	Arg	Leu	Asp	Phe	Pro	Ala	Leu	Ser	Ile	Glu	Asp	Gly	Ala	Asn	
115					120					125						
Ser	Ala	Ile	Glu	Val	Val	His	Ala	Leu	Gly	Ile	Lys	Gln	Leu	Ala	Ser	
130					135					140						
Leu	Ile	Gly	Asn	Ser	Met	Gly	Gly	Met	Thr	Ala	Leu	Ala	Ile	Leu	Leu	
145					150					155					160	
Leu	His	Pro	Asp	Ile	Ala	Arg	Ser	His	Ile	Asn	Ile	Ser	Gly	Ser	Ala	
165					170					175						
Gln	Ala	Leu	Pro	Phe	Ser	Ile	Ala	Ile	Arg	Ser	Leu	Gln	Arg	Glu	Ala	
180					185					190						
Ile	Arg	Leu	Asp	Pro	His	Trp	Arg	Gln	Gly	Asp	Tyr	Asp	Asp	Thr	His	
195					200					205						
Tyr	Pro	Glu	Ser	Gly	Leu	Arg	Ile	Ala	Arg	Lys	Leu	Gly	Val	Ile	Thr	
210					215					220						
Tyr	Arg	Ser	Ala	Leu	Glu	Trp	Asp	Gly	Arg	Phe	Gly	Arg	Val	Arg	Leu	
225					230					235					240	
Asp	Ser	Asp	Gln	Thr	Asn	Asp	Thr	Pro	Phe	Gly	Leu	Glu	Phe	Gln	Ile	
245					250					255						
Glu	Asn	Tyr	Leu	Glu	Ser	His	Ala	His	Arg	Phe	Val	His	Thr	Phe	Asp	
260					265					270						
Pro	Asn	Cys	Tyr	Leu	Tyr	Leu	Ser	Arg	Ser	Met	Asp	Trp	Phe	Asp	Val	
275					280					285						
Ala	Glu	Tyr	Ala	Asn	Gly	Asp	Ile	Leu	Ala	Gly	Leu	Ala	Arg	Ile	Arg	
290					295					300						
Ile	Gln	Arg	Ala	Leu	Ala	Ile	Gly	Ser	His	Thr	Asp	Ile	Leu	Phe	Pro	
305					310					315					320	
Ile	Gln	Gln	Gln	Gln	Gln	Ile	Ala	Glu	Gly	Leu	Arg	Arg	Gly	Gly	Thr	
325					330					335						
His	Ala	Thr	Phe	Leu	Gly	Leu	Asp	Ser	Pro	Gln	Gly	His	Asp	Ala	Phe	
340					345					350						
Leu	Val	Asp	Ile	Ala	Arg	Phe	Gly	Pro	Pro	Val	Lys	Glu	Phe	Leu	Asp	

355

360

365

Glu Leu
370

<210> 37

<211> 1113

<212> DNA

<213> Xylella oleander

<220>

<221> CDS

<222> (1)..(1110)

<223> RXYF01729

<400> 37

atg acc gaa ttt atc cct ccg ggc agc cta ttc cat gcg ctc tcc tct	48
Met Thr Glu Phe Ile Pro Pro Gly Ser Leu Phe His Ala Leu Ser Ser	
1 5 10 15	
cca ttt gcg atg aag cgt ggc gga caa ctc cac cac gcc cgc atc gct	96
Pro Phe Ala Met Lys Arg Gly Gly Gln Leu His His Ala Arg Ile Ala	
20 25 30	
tac gaa aca tgg ggc cgc ctc aat gcc agc gcc acc aat gcc att ctg	144
Tyr Glu Thr Trp Gly Arg Leu Asn Ala Ser Ala Thr Asn Ala Ile Leu	
35 40 45	
atc atg cct ggc tta tca ccc aat gca cat gcc gca cac cat gac agc	192
Ile Met Pro Gly Leu Ser Pro Asn Ala His Ala Ala His His Asp Ser	
50 55 60	
aat gct gag cca ggc tgg tgg gag tca atg cta ggt cca ggc aaa ccc	240
Asn Ala Glu Pro Gly Trp Trp Glu Ser Met Leu Gly Pro Gly Lys Pro	
65 70 75 80	
atc gac aca gac cgt tgg ttc gtg atc tgt gtc aac tca ctt ggt agc	288
Ile Asp Thr Asp Arg Trp Phe Val Ile Cys Val Asn Ser Leu Gly Ser	
85 90 95	
tgc aaa gga tcg act ggc cct gca tcg tac aac ccc atc acg cag gcc	336
Cys Lys Gly Ser Thr Gly Pro Ala Ser Tyr Asn Pro Ile Thr Gln Ala	
100 105 110	
atg tat cgt ttg gac ttt cca gca ctg tca atc gaa gac ggg gcc aac	384
Met Tyr Arg Leu Asp Phe Pro Ala Leu Ser Ile Glu Asp Gly Ala Asn	
115 120 125	
gcc gca att gaa gtg gta cat gca ctg ggc atc aag caa ctt gcc agc	432
Ala Ala Ile Glu Val Val His Ala Leu Gly Ile Lys Gln Leu Ala Ser	
130 135 140	
ctg atc ggc aat tca atg ggg ggc atg acg aca ctg gcc atc ctg ctg	480
Leu Ile Gly Asn Ser Met Gly Gly Met Thr Thr Leu Ala Ile Leu Leu	
145 150 155 160	
tta cat cca gat att gcc cgc agc cac atc aac atc tca ggc agc gcg	528
Leu His Pro Asp Ile Ala Arg Ser His Ile Asn Ile Ser Gly Ser Ala	
165 170 175	
cag gca tta ccg ttt tcc atc gcc att cgc tcg cta caa cgc gag gcg	576
Gln Ala Leu Pro Phe Ser Ile Ala Ile Arg Ser Leu Gln Arg Glu Ala	

180										185										190										
atc	cgc	ctg	gac	ccc	cat	tgg	aag	cag	gga	gac	tac	gac	gac	acc	cac	624														
Ile	Arg	Leu	Asp	Pro	His	Trp	Lys	Gln	Gly	Asp	Tyr	Asp	Asp	Thr	His															
195					200					205																				
tac	ccg	gaa	tcg	ggg	cta	cgc	atc	gca	cgc	aaa	ctc	ggg	gtg	atc	acc	672														
Tyr	Pro	Glu	Ser	Gly	Leu	Arg	Ile	Ala	Arg	Lys	Leu	Gly	Val	Ile	Thr															
210					215					220																				
tac	cgc	tcc	gcg	ctg	gaa	tgg	gac	ggg	cgt	ttt	ggc	cgg	gta	cgc	ttg	720														
Tyr	Arg	Ser	Ala	Leu	Glu	Trp	Asp	Gly	Arg	Phe	Gly	Arg	Val	Arg	Leu															
225					230					235					240															
gat	tcg	gac	caa	acc	aac	gac	aca	cca	ttc	gga	ctg	gaa	ttc	caa	att	768														
Asp	Ser	Asp	Gln	Thr	Asn	Asp	Thr	Pro	Phe	Gly	Leu	Glu	Phe	Gln	Ile															
245					250					255																				
gaa	aac	tac	ttg	gaa	agc	cat	gca	cac	cgc	ttc	gtg	cac	acc	ttc	gac	816														
Glu	Asn	Tyr	Leu	Glu	Ser	His	Ala	His	Arg	Phe	Val	His	Thr	Phe	Asp															
260					265					270																				
cca	aac	tgc	tac	ctg	tac	ctg	agc	cgc	tcc	atg	gac	tgg	ttc	gac	gtg	864														
Pro	Asn	Cys	Tyr	Leu	Tyr	Leu	Ser	Arg	Ser	Met	Asp	Trp	Phe	Asp	Val															
275					280					285																				
gcc	gag	tac	gcc	aat	gga	gac	att	ctt	gcc	ggg	ctg	gcc	agg	atc	cga	912														
Ala	Glu	Tyr	Ala	Asn	Gly	Asp	Ile	Leu	Ala	Gly	Leu	Ala	Arg	Ile	Arg															
290					295					300																				
atc	caa	cgc	gca	ctt	gcc	atc	ggg	agc	cat	acc	gac	atc	ctc	ttt	cca	960														
Ile	Gln	Arg	Ala	Leu	Ala	Ile	Gly	Ser	His	Thr	Asp	Ile	Leu	Phe	Pro															
305					310					315					320															
ata	caa	cag	caa	caa	caa	att	gcc	gaa	ggg	cta	cgc	cgt	ggc	ggg	aca	1008														
Ile	Gln	Gln	Gln	Gln	Gln	Ile	Ala	Glu	Gly	Leu	Arg	Arg	Gly	Gly	Thr															
325					330					335																				
cac	gcc	acc	ttc	ctg	ggc	ctt	gac	tca	ccg	cag	gga	cat	gat	gcg	ttc	1056														
His	Ala	Thr	Phe	Leu	Gly	Leu	Asp	Ser	Pro	Gln	Gly	His	Asp	Ala	Phe															
340					345					350																				
ctt	gtg	gat	atc	gca	gga	ttt	ggc	cct	cca	gtg	aag	gaa	ttt	ctg	ggc	1104														
Leu	Val	Asp	Ile	Ala	Gly	Phe	Gly	Pro	Pro	Val	Lys	Glu	Phe	Leu	Gly															
355					360					365																				
gaa	ctg	tga														1113														
Glu	Leu																													
370																														

<210> 38

<211> 370

<212> PRT

<213> Xylella oleander

<400> 38

Met	Thr	Glu	Phe	Ile	Pro	Pro	Gly	Ser	Leu	Phe	His	Ala	Leu	Ser	Ser
1					5				10					15	

Pro	Phe	Ala	Met	Lys	Arg	Gly	Gly	Gln	Leu	His	His	Ala	Arg	Ile	Ala
			20					25					30		

Tyr Glu Thr Trp Gly Arg Leu Asn Ala Ser Ala Thr Asn Ala Ile Leu
 35 40 45
 Ile Met Pro Gly Leu Ser Pro Asn Ala His Ala Ala His His Asp Ser
 50 55 60
 Asn Ala Glu Pro Gly Trp Trp Glu Ser Met Leu Gly Pro Gly Lys Pro
 65 70 75 80
 Ile Asp Thr Asp Arg Trp Phe Val Ile Cys Val Asn Ser Leu Gly Ser
 85 90 95
 Cys Lys Gly Ser Thr Gly Pro Ala Ser Tyr Asn Pro Ile Thr Gln Ala
 100 105 110
 Met Tyr Arg Leu Asp Phe Pro Ala Leu Ser Ile Glu Asp Gly Ala Asn
 115 120 125
 Ala Ala Ile Glu Val Val His Ala Leu Gly Ile Lys Gln Leu Ala Ser
 130 135 140
 Leu Ile Gly Asn Ser Met Gly Gly Met Thr Thr Leu Ala Ile Leu Leu
 145 150 155 160
 Leu His Pro Asp Ile Ala Arg Ser His Ile Asn Ile Ser Gly Ser Ala
 165 170 175
 Gln Ala Leu Pro Phe Ser Ile Ala Ile Arg Ser Leu Gln Arg Glu Ala
 180 185 190
 Ile Arg Leu Asp Pro His Trp Lys Gln Gly Asp Tyr Asp Asp Thr His
 195 200 205
 Tyr Pro Glu Ser Gly Leu Arg Ile Ala Arg Lys Leu Gly Val Ile Thr
 210 215 220
 Tyr Arg Ser Ala Leu Glu Trp Asp Gly Arg Phe Gly Arg Val Arg Leu
 225 230 235 240
 Asp Ser Asp Gln Thr Asn Asp Thr Pro Phe Gly Leu Glu Phe Gln Ile
 245 250 255
 Glu Asn Tyr Leu Glu Ser His Ala His Arg Phe Val His Thr Phe Asp
 260 265 270
 Pro Asn Cys Tyr Leu Tyr Leu Ser Arg Ser Met Asp Trp Phe Asp Val
 275 280 285
 Ala Glu Tyr Ala Asn Gly Asp Ile Leu Ala Gly Leu Ala Arg Ile Arg
 290 295 300
 Ile Gln Arg Ala Leu Ala Ile Gly Ser His Thr Asp Ile Leu Phe Pro
 305 310 315 320
 Ile Gln Gln Gln Gln Gln Ile Ala Glu Gly Leu Arg Arg Gly Gly Thr
 325 330 335
 His Ala Thr Phe Leu Gly Leu Asp Ser Pro Gln Gly His Asp Ala Phe
 340 345 350
 Leu Val Asp Ile Ala Gly Phe Gly Pro Pro Val Lys Glu Phe Leu Gly
 355 360 365

Glu Leu
370

<210> 39

<211> 1578

<212> DNA

<213> Emericella nidulans

<220>

<221> CDS

<222> (1) .. (1575)

<223> REN00010

<400> 39

atg agt ccg ctg aac ggc gtc gct cgt tcc ttt ccg cgg ccc ttc cag	48
Met Ser Pro Leu Asn Gly Val Ala Arg Ser Phe Pro Arg Pro Phe Gln	
1 5 10 15	
gcc gtg acc agg cgg cct ttt cga gtt gtc cag ccg gcc atc gcc tgt	96
Ala Val Thr Arg Arg Pro Phe Arg Val Val Gln Pro Ala Ile Ala Cys	
20 25 30	
ccg tcc aac agc cgg tcg ttt aac cat tct cga tca tta cga tca acg	144
Pro Ser Asn Ser Arg Ser Phe Asn His Ser Arg Ser Leu Arg Ser Thr	
35 40 45	
ggg tct cag tcc ccc gct cca tcc cca cgc gac tcc tcg aat ccc gcg	192
Gly Ser Gln Ser Pro Ala Pro Ser Pro Arg Asp Ser Ser Asn Pro Ala	
50 55 60	
ctg tcc ttc cct tgc ctc gac gcc cag gag gcc aag tcc gct ctt ctt	240
Leu Ser Phe Pro Cys Leu Asp Ala Gln Glu Ala Lys Ser Ala Leu Leu	
65 70 75 80	
tcc gcg cga tct ctt ggt tca ggc cct gaa ccc tcc tat acc gcc ggc	288
Ser Ala Arg Ser Leu Gly Ser Gly Pro Glu Pro Ser Tyr Thr Ala Gly	
85 90 95	
cac cac gaa cga ttc cat tcc gac gaa ccg ctg ctc ctt gat tgg ggc	336
His His Glu Arg Phe His Ser Asp Glu Pro Leu Leu Leu Asp Trp Gly	
100 105 110	
ggt ttg ctt cca gaa ttt gat atc gca tat gag aca tgg ggc cag ctg	384
Gly Leu Leu Pro Glu Phe Asp Ile Ala Tyr Glu Thr Trp Gly Gln Leu	
115 120 125	
aac gag aag aag gat aat gtc att ctg ctg cat acc ggt ctg tct gca	432
Asn Glu Lys Lys Asp Asn Val Ile Leu Leu His Thr Gly Leu Ser Ala	
130 135 140	
tct agc cat gcg cac agc acc gaa gcg aac ccg aag ccc ggc tgg tgg	480
Ser Ser His Ala His Ser Thr Glu Ala Asn Pro Lys Pro Gly Trp Trp	
145 150 155 160	
gag aaa ttc ata ggt cct ggg aag acg cta gat acg gac aag tac ttt	528
Glu Lys Phe Ile Gly Pro Gly Lys Thr Leu Asp Thr Asp Lys Tyr Phe	
165 170 175	
gtg atc tgc acc aat gtc ctt gga ggg tgc tac ggt agc acg ggg ccc	576
Val Ile Cys Thr Asn Val Leu Gly Gly Cys Tyr Gly Ser Thr Gly Pro	
180 185 190	

tcg acg gtg gac ccg tcg gat ggg aag aag tat gct acg cgg ttt ccc	624
Ser Thr Val Asp Pro Ser Asp Gly Lys Lys Tyr Ala Thr Arg Phe Pro	
195 200 205	
atc ctg aca att gaa gat atg gtg cga gcg cag ttc cgc ctt ttg gac	672
Ile Leu Thr Ile Glu Asp Met Val Arg Ala Gln Phe Arg Leu Leu Asp	
210 215 220	
cat ctt ggg gtt cgg aaa ctc tac gcg tcc gtc ggc tcc agc atg ggt	720
His Leu Gly Val Arg Lys Leu Tyr Ala Ser Val Gly Ser Ser Met Gly	
225 230 235 240	
ggt atg cag agt ctt gca gcc ggt gtt ctg ttc cca gag cga gtg ggc	768
Gly Met Gln Ser Leu Ala Ala Gly Val Leu Phe Pro Glu Arg Val Gly	
245 250 255	
aag att gtg tcg att agc ggt tgt gct cga agc cat ccg tac agc att	816
Lys Ile Val Ser Ile Ser Gly Cys Ala Arg Ser His Pro Tyr Ser Ile	
260 265 270	
gct atg cgc cat acc cag cgg cag gtg ttg atg atg gat cca aat tgg	864
Ala Met Arg His Thr Gln Arg Gln Val Leu Met Met Asp Pro Asn Trp	
275 280 285	
gct cga ggt ttc tac tac gat tcg atc cca cct cat tca ggc atg aag	912
Ala Arg Gly Phe Tyr Tyr Asp Ser Ile Pro Pro His Ser Gly Met Lys	
290 295 300	
ctc gct cgc gag att gcc acc gtc acg tac cgc agc gga cca gaa tgg	960
Leu Ala Arg Glu Ile Ala Thr Val Thr Arg Ser Gly Pro Glu Trp	
305 310 315 320	
gag aaa cgc ttt ggt cgg aaa cgg gct gat ccg agc aaa cag cct gcg	1008
Glu Lys Arg Phe Gly Arg Lys Arg Ala Asp Pro Ser Lys Gln Pro Ala	
325 330 335	
ctt tgc ccc gac ttt ctc atc gag acg tat ctc gac cac gcc ggt gaa	1056
Leu Cys Pro Asp Phe Leu Ile Glu Thr Tyr Leu Asp His Ala Gly Glu	
340 345 350	
aaa ttc tgc ttg gaa tac gat gcc aac agc ctg ctc tac atc tcc aag	1104
Lys Phe Cys Leu Glu Tyr Asp Ala Asn Ser Leu Leu Tyr Ile Ser Lys	
355 360 365	
gcg atg gat ctg ttt gac cta ggg ttg act cag caa ctc gcg acg aag	1152
Ala Met Asp Leu Phe Asp Leu Gly Leu Thr Gln Gln Leu Ala Thr Lys	
370 375 380	
aag cag agg gcg gag gcc cag gcg aag att agc agc gga aca aac act	1200
Lys Gln Arg Ala Glu Ala Gln Ala Lys Ile Ser Ser Gly Thr Asn Thr	
385 390 395 400	
gtc aat gat gcg tcg tgc agc ctt aca ctt cct gaa cag cca tac cag	1248
Val Asn Asp Ala Ser Cys Ser Leu Thr Leu Pro Glu Gln Pro Tyr Gln	
405 410 415	
gag cag cca tct gcc tcg aca tcc gcc gag cag tct gct tcc gct tca	1296
Glu Gln Pro Ser Ala Ser Thr Ser Ala Glu Gln Ser Ala Ser Ala Ser	
420 425 430	
gag acc ggg tcg gct ccg aac gat ctt gtt gcc ggg ctt gcg ccg ctg	1344
Glu Thr Gly Ser Ala Pro Asn Asp Leu Val Ala Gly Leu Ala Pro Leu	
435 440 445	

aaa gac cat cag gtg ctg gta atc gga gtc gca agc gac att ctc ttc 1392
 Lys Asp His Gln Val Leu Val Ile Gly Val Ala Ser Asp Ile Leu Phe
 450 455 460

ccg gcg tgg caa cag cgc gag atc gcg gag act ctg att caa gca ggg 1440
 Pro Ala Trp Gln Gln Arg Glu Ile Ala Glu Thr Leu Ile Gln Ala Gly
 465 470 475 480

aac aag acc gtg gag cat att gag ctg ggc aac gac gtg tct ctc ttt 1488
 Asn Lys Thr Val Glu His Ile Glu Leu Asn Asp Val Ser Leu Phe
 485 490 495

ggt cat gac aca ttc ctc ctt gat gtc aga acg tcg gag gcg cag ttc 1536
 Gly His Asp Thr Phe Leu Leu Asp Val Arg Thr Ser Glu Ala Gln Phe
 500 505 510

gca agt tcc gta cta gtc ggc tcg cac ata att gta caa tag 1578
 Ala Ser Ser Val Leu Val Gly Ser His Ile Ile Val Gln
 515 520 525

<210> 40

<211> 525

<212> PRT

<213> Emericella nidulans

<400> 40

Met Ser Pro Leu Asn Gly Val Ala Arg Ser Phe Pro Arg Pro Phe Gln
 1 5 10 15

Ala Val Thr Arg Arg Pro Phe Arg Val Val Gln Pro Ala Ile Ala Cys
 20 25 30

Pro Ser Asn Ser Arg Ser Phe Asn His Ser Arg Ser Leu Arg Ser Thr
 35 40 45

Gly Ser Gln Ser Pro Ala Pro Ser Pro Arg Asp Ser Ser Asn Pro Ala
 50 55 60

Leu Ser Phe Pro Cys Leu Asp Ala Gln Glu Ala Lys Ser Ala Leu Leu
 65 70 75 80

Ser Ala Arg Ser Leu Gly Ser Gly Pro Glu Pro Ser Tyr Thr Ala Gly
 85 90 95

His His Glu Arg Phe His Ser Asp Glu Pro Leu Leu Leu Asp Trp Gly
 100 105 110

Gly Leu Leu Pro Glu Phe Asp Ile Ala Tyr Glu Thr Trp Gly Gln Leu
 115 120 125

Asn Glu Lys Lys Asp Asn Val Ile Leu Leu His Thr Gly Leu Ser Ala
 130 135 140

Ser Ser His Ala His Ser Thr Glu Ala Asn Pro Lys Pro Gly Trp Trp
 145 150 155 160

Glu Lys Phe Ile Gly Pro Gly Lys Thr Leu Asp Thr Asp Lys Tyr Phe
 165 170 175

Val Ile Cys Thr Asn Val Leu Gly Gly Cys Tyr Gly Ser Thr Gly Pro
 180 185 190

Ser Thr Val Asp Pro Ser Asp Gly Lys Lys Tyr Ala Thr Arg Phe Pro
 195 200 205
 Ile Leu Thr Ile Glu Asp Met Val Arg Ala Gln Phe Arg Leu Leu Asp
 210 215 220
 His Leu Gly Val Arg Lys Leu Tyr Ala Ser Val Gly Ser Ser Met Gly
 225 230 235 240
 Gly Met Gln Ser Leu Ala Ala Gly Val Leu Phe Pro Glu Arg Val Gly
 245 250 255
 Lys Ile Val Ser Ile Ser Gly Cys Ala Arg Ser His Pro Tyr Ser Ile
 260 265 270
 Ala Met Arg His Thr Gln Arg Gln Val Leu Met Met Asp Pro Asn Trp
 275 280 285
 Ala Arg Gly Phe Tyr Tyr Asp Ser Ile Pro Pro His Ser Gly Met Lys
 290 295 300
 Leu Ala Arg Glu Ile Ala Thr Val Thr Tyr Arg Ser Gly Pro Glu Trp
 305 310 315 320
 Glu Lys Arg Phe Gly Arg Lys Arg Ala Asp Pro Ser Lys Gln Pro Ala
 325 330 335
 Leu Cys Pro Asp Phe Leu Ile Glu Thr Tyr Leu Asp His Ala Gly Glu
 340 345 350
 Lys Phe Cys Leu Glu Tyr Asp Ala Asn Ser Leu Leu Tyr Ile Ser Lys
 355 360 365
 Ala Met Asp Leu Phe Asp Leu Gly Leu Thr Gln Gln Leu Ala Thr Lys
 370 375 380
 Lys Gln Arg Ala Glu Ala Gln Ala Lys Ile Ser Ser Gly Thr Asn Thr
 385 390 395 400
 Val Asn Asp Ala Ser Cys Ser Leu Thr Leu Pro Glu Gln Pro Tyr Gln
 405 410 415
 Glu Gln Pro Ser Ala Ser Thr Ser Ala Glu Gln Ser Ala Ser Ala Ser
 420 425 430
 Glu Thr Gly Ser Ala Pro Asn Asp Leu Val Ala Gly Leu Ala Pro Leu
 435 440 445
 Lys Asp His Gln Val Leu Val Ile Gly Val Ala Ser Asp Ile Leu Phe
 450 455 460
 Pro Ala Trp Gln Gln Arg Glu Ile Ala Glu Thr Leu Ile Gln Ala Gly
 465 470 475 480
 Asn Lys Thr Val Glu His Ile Glu Leu Gly Asn Asp Val Ser Leu Phe
 485 490 495
 Gly His Asp Thr Phe Leu Leu Asp Val Arg Thr Ser Glu Ala Gln Phe
 500 505 510
 Ala Ser Ser Val Leu Val Gly Ser His Ile Ile Val Gln
 515 520 525

<210> 41
 <211> 1170
 <212> DNA
 <213> Mesorhizobium loti

<220>
 <221> CDS
 <222> (1)..(1167)
 <223> NP_104621

<400> 41
 atg gcc gct ctg cgc gca gga aag acc aac aac gag gcc gac cag ccg 48
 Met Ala Ala Leu Arg Ala Gly Lys Thr Asn Asn Glu Ala Asp Gln Pro
 1 5 10 15

tcg agc ccg gtg ttg cgc ttc ggg gcg gac aag ccg ctc aag ctc gac 96
 Ser Ser Pro Val Leu Arg Phe Gly Ala Asp Lys Pro Leu Lys Leu Asp
 20 25 30

gcc ggc acg ctt ttg tcg ccg ttc cag atc gcc tat cag acc tac ggc 144
 Ala Gly Thr Leu Leu Ser Pro Phe Gln Ile Ala Tyr Gln Thr Tyr Gly
 35 40 45

acg ctg aac gat gcc cgc tcc aat gcc atc ctc gtc tgc cat gcg ctg 192
 Thr Leu Asn Asp Ala Arg Ser Asn Ala Ile Leu Val Cys His Ala Leu
 50 55 60

acc ggc gac cag cat gtc gcc aac acc aat ccg gtg acc ggc aag ccg 240
 Thr Gly Asp Gln His Val Ala Asn Thr Asn Pro Val Thr Gly Lys Pro
 65 70 75 80

gga tgg tgg gaa gtg ctg atc ggc ccc ggc agg atc atc gac acc aac 288
 Gly Trp Trp Glu Val Leu Ile Gly Pro Gly Arg Ile Ile Asp Thr Asn
 85 90 95

cgt ttc ttc gtc atc tgc tcc aac gtc atc ggc ggt tgt ctg ggc tcc 336
 Arg Phe Phe Val Ile Cys Ser Asn Val Ile Gly Gly Cys Leu Gly Ser
 100 105 110

acc ggc ccg gcc tcg acc aac ccc gcc acc ggc aag ccc tac ggg ctc 384
 Thr Gly Pro Ala Ser Thr Asn Pro Ala Thr Gly Lys Pro Tyr Gly Leu
 115 120 125

gac ctg ccg gtc atc acc atc cgc gat atg gtg cgc gcg cag cag atg 432
 Asp Leu Pro Val Ile Thr Ile Arg Asp Met Val Arg Ala Gln Gln Met
 130 135 140

ctg atc gat cat ttc ggc atc gag aaa ctg ttc tgc gtg ctc ggc ggc 480
 Leu Ile Asp His Phe Gly Ile Glu Lys Leu Phe Cys Val Leu Gly Gly
 145 150 155 160

tcg atg ggc gga atg cag gtg ctg gaa tgg gcg tcg agc tac ccc gag 528
 Ser Met Gly Gly Met Gln Val Leu Glu Trp Ala Ser Ser Tyr Pro Glu
 165 170 175

cgc gtc ttt tcg gca ctg ccg atc gcc acc ggc gcg cgc cat tcc tcg 576
 Arg Val Phe Ser Ala Leu Pro Ile Ala Thr Gly Ala Arg His Ser Ser
 180 185 190

cag aac atc gcc ttc cac gag gtc ggc cgg cag gct gtc atg gcc gat 624
 Gln Asn Ile Ala Phe His Glu Val Gly Arg Gln Ala Val Met Ala Asp

195	200	205	
ccg gac tgg cac ggc ggc aaa tat ttc gaa aac ggc aaa cgc ccg gaa Pro Asp Trp His Gly Gly Lys Tyr Phe Glu Asn Gly Lys Arg Pro Glu 210 215 220			672
aag ggc ctg gcg gta gcg cgc atg gcc gcc cac ata acc tat ctg tcg Lys Gly Leu Ala Val Ala Arg Met Ala Ala His Ile Thr Tyr Leu Ser 225 230 235 240			720
gaa gcc gcc ctg cac cgg aaa ttc ggc cgc aat ctg cag gat cgc gag Glu Ala Ala Leu His Arg Lys Phe Gly Arg Asn Leu Gln Asp Arg Glu 245 250 255			768
gcg ctg acc ttc ggc ttc gac gcc gac ttc cag atc gaa agc tat ctg Ala Leu Thr Phe Gly Phe Asp Ala Asp Phe Gln Ile Glu Ser Tyr Leu 260 265 270			816
cgc cac caa ggc atg acc ttc gtc gac cgc ttc gac gcc aat tcc tat Arg His Gln Gly Met Thr Phe Val Asp Arg Phe Asp Ala Asn Ser Tyr 275 280 285			864
ctc tac atg acg cgg tcg atg gac tat ttc gac ctc gcc gcc gat cat Leu Tyr Met Thr Arg Ser Met Asp Tyr Phe Asp Leu Ala Ala Asp His 290 295 300			912
ggc ggg cgg ctg gcg gat gcc ttt gcc ggc acc aaa acc cgc ttc tgc Gly Gly Arg Leu Ala Asp Ala Phe Ala Gly Thr Lys Thr Arg Phe Cys 305 310 315 320			960
ctg gtg tcc ttc acc tcg gat tgg ttg ttt ccg acc gaa gag agc cgc Leu Val Ser Phe Thr Ser Asp Trp Leu Phe Pro Thr Glu Glu Ser Arg 325 330 335			1008
tcg atc gtg cac gcg ctc aac gcc gcc ggc gcg tcc gtg tcc ttc gtc Ser Ile Val His Ala Leu Asn Ala Ala Gly Ala Ser Val Ser Phe Val 340 345 350			1056
gaa atc gag acc gac cgc ggc cac gat gcc ttc ctg ctc gac gag ccg Glu Ile Glu Thr Asp Arg Gly His Asp Ala Phe Leu Leu Asp Glu Pro 355 360 365			1104
gaa ctg ttc gcc gcc atc aac ggc ttc atc ggc tcc gcg gcg cgg gcg Glu Leu Phe Ala Ala Ile Asn Gly Phe Ile Gly Ser Ala Ala Arg Ala 370 375 380			1152
aga ggg cta agc gca tga Arg Gly Leu Ser Ala 385			1170

<210> 42

<211> 389

<212> PRT

<213> Mesorhizobium loti

<400> 42

Met	Ala	Ala	Leu	Arg	Ala	Gly	Lys	Thr	Asn	Asn	Glu	Ala	Asp	Gln	Pro
1					5				10					15	
Ser	Ser	Pro	Val	Leu	Arg	Phe	Gly	Ala	Asp	Lys	Pro	Leu	Lys	Leu	Asp
			20					25					30		

Ala Gly Thr Leu Leu Ser Pro Phe Gln Ile Ala Tyr Gln Thr Tyr Gly
 35 40 45
 Thr Leu Asn Asp Ala Arg Ser Asn Ala Ile Leu Val Cys His Ala Leu
 50 55 60
 Thr Gly Asp Gln His Val Ala Asn Thr Asn Pro Val Thr Gly Lys Pro
 65 70 75 80
 Gly Trp Trp Glu Val Leu Ile Gly Pro Gly Arg Ile Ile Asp Thr Asn
 85 90 95
 Arg Phe Phe Val Ile Cys Ser Asn Val Ile Gly Gly Cys Leu Gly Ser
 100 105 110
 Thr Gly Pro Ala Ser Thr Asn Pro Ala Thr Gly Lys Pro Tyr Gly Leu
 115 120 125
 Asp Leu Pro Val Ile Thr Ile Arg Asp Met Val Arg Ala Gln Gln Met
 130 135 140
 Leu Ile Asp His Phe Gly Ile Glu Lys Leu Phe Cys Val Leu Gly Gly
 145 150 155 160
 Ser Met Gly Gly Met Gln Val Leu Glu Trp Ala Ser Ser Tyr Pro Glu
 165 170 175
 Arg Val Phe Ser Ala Leu Pro Ile Ala Thr Gly Ala Arg His Ser Ser
 180 185 190
 Gln Asn Ile Ala Phe His Glu Val Gly Arg Gln Ala Val Met Ala Asp
 195 200 205
 Pro Asp Trp His Gly Gly Lys Tyr Phe Glu Asn Gly Lys Arg Pro Glu
 210 215 220
 Lys Gly Leu Ala Val Ala Arg Met Ala Ala His Ile Thr Tyr Leu Ser
 225 230 235 240
 Glu Ala Ala Leu His Arg Lys Phe Gly Arg Asn Leu Gln Asp Arg Glu
 245 250 255
 Ala Leu Thr Phe Gly Phe Asp Ala Asp Phe Gln Ile Glu Ser Tyr Leu
 260 265 270
 Arg His Gln Gly Met Thr Phe Val Asp Arg Phe Asp Ala Asn Ser Tyr
 275 280 285
 Leu Tyr Met Thr Arg Ser Met Asp Tyr Phe Asp Leu Ala Ala Asp His
 290 295 300
 Gly Gly Arg Leu Ala Asp Ala Phe Ala Gly Thr Lys Thr Arg Phe Cys
 305 310 315 320
 Leu Val Ser Phe Thr Ser Asp Trp Leu Phe Pro Thr Glu Glu Ser Arg
 325 330 335
 Ser Ile Val His Ala Leu Asn Ala Ala Gly Ala Ser Val Ser Phe Val
 340 345 350
 Glu Ile Glu Thr Asp Arg Gly His Asp Ala Phe Leu Leu Asp Glu Pro
 355 360 365

Glu Leu Phe Ala Ala Ile Asn Gly Phe Ile Gly Ser Ala Ala Arg Ala
 370 375 380

Arg Gly Leu Ser Ala
 385

<210> 43

<211> 1155

<212> DNA

<213> acremonium crysogenum

<220>

<221> CDS

<222> (1) .. (1152)

<223> P39058

<400> 43

tgt cgc ctc aga tcg cca atc gct tcg agg ctt cgc tag atg ccc aag	48
Cys Arg Leu Arg Ser Pro Ile Ala Ser Arg Leu Arg Xaa Met Pro Lys	
1 5 10 15	
aca tag cca gaa tat cgc tct tca cac tgg aat ctg gcg tca tcc ttc	96
Thr Xaa Pro Glu Tyr Arg Ser Ser His Trp Asn Leu Ala Ser Ser Phe	
20 25 30	
gcg atg tac ccg tgg cat aca aat cgt ggg gtc gca tga atg tct caa	144
Ala Met Tyr Pro Trp His Thr Asn Arg Gly Val Ala Xaa Met Ser Gln	
35 40 45	
ggg ata act gcg tca tcg tct gcc aca cct tga cga gca gcg ccc atg	192
Gly Ile Thr Ala Ser Ser Ser Ala Thr Pro Xaa Arg Ala Ala Pro Met	
50 55 60	
tca cct cgt ggt ggc cca cac tgt ttg gcc aag gca ggg ctt tcg ata	240
Ser Pro Arg Gly Gly Pro His Cys Leu Ala Lys Ala Gly Leu Ser Ile	
65 70 75 80	
cct ctc gct act tca tca tct gcc taa att atc tcg gga gcc cct ttg	288
Pro Leu Ala Thr Ser Ser Ser Ala Xaa Ile Ile Ser Gly Ala Pro Leu	
85 90 95	
gga gtg ctg gac cat gtt cac cgg acc ccg atg cag aag gcc agc gcc	336
Gly Val Leu Asp His Val His Arg Thr Pro Met Gln Lys Ala Ser Ala	
100 105 110	
cgt acg ggg cca agt ttc ctc gca cga cga ttc gag atg atg ttc gta	384
Arg Thr Gly Pro Ser Phe Leu Ala Arg Arg Phe Glu Met Met Phe Val	
115 120 125	
ttc atc gcc agg tgc tcg aca ggt tag gcg tca ggc aaa ttg ctg ccg	432
Phe Ile Ala Arg Cys Ser Thr Gly Xaa Ala Ser Gly Lys Leu Leu Pro	
130 135 140	
tag tcg gcg cat cca tgg gtg gaa tgc aca ctc tgg aat ggg cct tct	480
Xaa Ser Ala His Pro Trp Val Glu Cys Thr Leu Trp Asn Gly Pro Ser	
145 150 155 160	
ttg gtc ccg agt acg tgc gaa aga ttg tgc cca tcg cga cat cat gcc	528
Leu Val Pro Ser Thr Cys Glu Arg Leu Cys Pro Ser Arg His His Ala	
165 170 175	
gtc aga gcg gct ggt gcg cag ctt ggt tcg aga cac aga ggc agt gca	576

Val	Arg	Ala	Ala	Gly	Ala	Gln	Leu	Gly	Ser	Arg	His	Arg	Gly	Ser	Ala		
			180					185						190			
tct	atg	atg	acc	cca	agt	acc	tgg	acg	ggg	agt	acg	acg	tag	acg	acc	624	
Ser	Met	Met	Thr	Pro	Ser	Thr	Trp	Thr	Gly	Ser	Thr	Thr	Xaa	Thr	Thr		
		195					200					205					
agc	ctg	tcc	ggg	ggc	tcg	aaa	cag	cgc	gca	aga	ttg	cga	atc	tca	cgt	672	
Ser	Leu	Ser	Gly	Gly	Ser	Lys	Gln	Arg	Ala	Arg	Leu	Arg	Ile	Ser	Arg		
		210				215					220						
aca	aga	gca	aac	ctg	cga	tgg	acg	agc	gct	tcc	ata	tgg	ctc	cag	gag	720	
Thr	Arg	Ala	Asn	Leu	Arg	Trp	Thr	Ser	Ala	Ser	Ile	Trp	Leu	Gln	Glu		
225					230				235						240		
tcc	aag	ccg	gcc	gga	ata	tca	gca	gcc	agg	atg	cga	aga	agg	aaa	tca	768	
Ser	Lys	Pro	Ala	Gly	Ile	Ser	Ala	Ala	Arg	Met	Arg	Arg	Arg	Lys	Ser		
			245						250					255			
acg	gca	cag	aca	gcg	gca	aca	gcc	acc	gtg	ctg	gcc	agc	cca	ttg	aag	816	
Thr	Ala	Gln	Thr	Ala	Ala	Thr	Ala	Thr	Val	Leu	Ala	Ser	Pro	Leu	Lys		
			260					265						270			
ccg	tat	ctt	cct	atc	tcc	ggc	acc	agg	ccc	aga	agt	ttg	ccg	cga	gct	864	
Pro	Tyr	Leu	Pro	Ile	Ser	Gly	Thr	Arg	Pro	Arg	Ser	Leu	Pro	Arg	Ala		
		275					280					285					
tcg	acg	cca	act	gct	aca	tcg	cca	tga	cac	tca	agt	tcg	aca	ccc	acg	912	
Ser	Thr	Pro	Thr	Ala	Thr	Ser	Pro	Xaa	His	Ser	Ser	Ser	Thr	Pro	Thr		
		290				295					300						
aca	tca	gca	gag	gcc	ggg	cag	gat	caa	tcc	cgg	agg	ctc	tgg	caa	tga	960	
Thr	Ser	Ala	Glu	Ala	Gly	Gln	Asp	Gln	Ser	Arg	Arg	Leu	Trp	Gln	Xaa		
305					310					315					320		
tta	cac	aac	cag	cgt	tga	tca	ttt	gcg	cca	ggc	cag	acg	gtc	tgt	act	1008	
Leu	His	Asn	Gln	Arg	Xaa	Ser	Phe	Ala	Pro	Gly	Gln	Thr	Val	Cys	Thr		
			325						330					335			
cgt	ttg	acg	agc	acg	ttg	aga	tgg	ggc	gca	gta	tcc	caa	aca	gtc	gtc	1056	
Arg	Leu	Thr	Ser	Thr	Leu	Arg	Trp	Gly	Ala	Val	Ser	Gln	Thr	Val	Val		
			340					345					350				
ttt	gcg	tgg	tgg	aca	cga	atg	agg	gtc	atg	act	tct	ttg	taa	tgg	aag	1104	
Phe	Ala	Trp	Trp	Thr	Arg	Met	Arg	Val	Met	Thr	Ser	Leu	Xaa	Trp	Lys		
		355					360					365					
cgg	aca	agg	tta	atg	atg	ccg	tca	gag	gat	tcc	tcg	atc	agt	cat	taa	1152	
Arg	Thr	Arg	Leu	Met	Met	Pro	Ser	Glu	Asp	Ser	Ser	Ile	Ser	His	Xaa		
		370				375					380						
tgt																1155	

<210> 44

<211> 384

<212> PRT

<213> acremonium crysogenum

<220>

<221> unsure

<222> 13 .. 13

<223> All occurrences of Xaa indicate any amino acid

<220>
<221> unsure
<222> 18 .. 18
<223> All occurrences of Xaa indicate any amino acid

<220>
<221> unsure
<222> 45 .. 45
<223> All occurrences of Xaa indicate any amino acid

<220>
<221> unsure
<222> 59 .. 59
<223> All occurrences of Xaa indicate any amino acid

<220>
<221> unsure
<222> 89 .. 89
<223> All occurrences of Xaa indicate any amino acid

<220>
<221> unsure
<222> 137 .. 137
<223> All occurrences of Xaa indicate any amino acid

<220>
<221> unsure
<222> 145 .. 145
<223> All occurrences of Xaa indicate any amino acid

<220>
<221> unsure
<222> 206 .. 206
<223> All occurrences of Xaa indicate any amino acid

<220>
<221> unsure
<222> 297 .. 297
<223> All occurrences of Xaa indicate any amino acid

<220>
<221> unsure
<222> 320 .. 320
<223> All occurrences of Xaa indicate any amino acid

<220>
<221> unsure
<222> 326 .. 326
<223> All occurrences of Xaa indicate any amino acid

<220>
<221> unsure
<222> 366 .. 366
<223> All occurrences of Xaa indicate any amino acid

<220>
<221> unsure
<222> 384 .. 384
<223> All occurrences of Xaa indicate any amino acid

<400> 44
Cys Arg Leu Arg Ser Pro Ile Ala Ser Arg Leu Arg Xaa Met Pro Lys

1	5	10	15
Thr Xaa Pro Glu Tyr Arg Ser Ser His Trp Asn Leu Ala Ser Ser Phe	20	25	30
Ala Met Tyr Pro Trp His Thr Asn Arg Gly Val Ala Xaa Met Ser Gln	35	40	45
Gly Ile Thr Ala Ser Ser Ser Ala Thr Pro Xaa Arg Ala Ala Pro Met	50	55	60
Ser Pro Arg Gly Gly Pro His Cys Leu Ala Lys Ala Gly Leu Ser Ile	65	70	75
Pro Leu Ala Thr Ser Ser Ser Ala Xaa Ile Ile Ser Gly Ala Pro Leu	85	90	95
Gly Val Leu Asp His Val His Arg Thr Pro Met Gln Lys Ala Ser Ala	100	105	110
Arg Thr Gly Pro Ser Phe Leu Ala Arg Arg Phe Glu Met Met Phe Val	115	120	125
Phe Ile Ala Arg Cys Ser Thr Gly Xaa Ala Ser Gly Lys Leu Leu Pro	130	135	140
Xaa Ser Ala His Pro Trp Val Glu Cys Thr Leu Trp Asn Gly Pro Ser	145	150	155
Leu Val Pro Ser Thr Cys Glu Arg Leu Cys Pro Ser Arg His His Ala	165	170	175
Val Arg Ala Ala Gly Ala Gln Leu Gly Ser Arg His Arg Gly Ser Ala	180	185	190
Ser Met Met Thr Pro Ser Thr Trp Thr Gly Ser Thr Thr Xaa Thr Thr	195	200	205
Ser Leu Ser Gly Gly Ser Lys Gln Arg Ala Arg Leu Arg Ile Ser Arg	210	215	220
Thr Arg Ala Asn Leu Arg Trp Thr Ser Ala Ser Ile Trp Leu Gln Glu	225	230	235
Ser Lys Pro Ala Gly Ile Ser Ala Ala Arg Met Arg Arg Arg Lys Ser	245	250	255
Thr Ala Gln Thr Ala Ala Thr Ala Thr Val Leu Ala Ser Pro Leu Lys	260	265	270
Pro Tyr Leu Pro Ile Ser Gly Thr Arg Pro Arg Ser Leu Pro Arg Ala	275	280	285
Ser Thr Pro Thr Ala Thr Ser Pro Xaa His Ser Ser Ser Thr Pro Thr	290	295	300
Thr Ser Ala Glu Ala Gly Gln Asp Gln Ser Arg Arg Leu Trp Gln Xaa	305	310	315
Leu His Asn Gln Arg Xaa Ser Phe Ala Pro Gly Gln Thr Val Cys Thr	325	330	335
Arg Leu Thr Ser Thr Leu Arg Trp Gly Ala Val Ser Gln Thr Val Val			

340

345

350

Phe Ala Trp Trp Thr Arg Met Arg Val Met Thr Ser Leu Xaa Trp Lys
 355 360 365

Arg Thr Arg Leu Met Met Pro Ser Glu Asp Ser Ser Ile Ser His Xaa
 370 375 380

<210> 45

<211> 1077

<212> DNA

<213> *Pseudomonas putida*

<220>

<221> CDS

<222> (1) .. (1074)

<223> AAK49778

<400> 45

atg tca act gtc ttt ccc gaa gat tcc gtc ggt ctg gta gta cgg caa 48
 Met Ser Thr Val Phe Pro Glu Asp Ser Val Gly Leu Val Val Arg Gln
 1 5 10 15

acc tcc cgg ttc gat gaa ccg ctg gca ctg gcc tgt ggc cgt tca ctg 96
 Thr Ser Arg Phe Asp Glu Pro Leu Ala Leu Ala Cys Gly Arg Ser Leu
 20 25 30

gcc agt tac gaa ctg gtc tac gag acc tat ggc acc ctg aac gcc agc 144
 Ala Ser Tyr Glu Leu Val Tyr Glu Thr Tyr Gly Thr Leu Asn Ala Ser
 35 40 45

gcg agc aac gcc gtg ctg atc tgc cat gcc ctg tcc ggc cac cac cat 192
 Ala Ser Asn Ala Val Leu Ile Cys His Ala Leu Ser Gly His His His
 50 55 60

gcc gct ggc tac cat gcc gcc acc gac cgc aag ccg ggc tgg tgg gac 240
 Ala Ala Gly Tyr His Ala Ala Thr Asp Arg Lys Pro Gly Trp Trp Asp
 65 70 75 80

agc tgc atc ggc ccc gga aaa ccg atc gat acc aac cgc ttc ttc gtg 288
 Ser Cys Ile Gly Pro Gly Lys Pro Ile Asp Thr Asn Arg Phe Phe Val
 85 90 95

gtc agc ctg aac aac ctc ggc ggc tgc aac ggc agc acc ggc ccc agc 336
 Val Ser Leu Asn Asn Leu Gly Gly Cys Asn Gly Ser Thr Gly Pro Ser
 100 105 110

agt gtc aac cca gcc acc ggt aaa ccc tat ggc gcc gag ttc ccg gta 384
 Ser Val Asn Pro Ala Thr Gly Lys Pro Tyr Gly Ala Glu Phe Pro Val
 115 120 125

ttg acc gtg gaa gac tgg gtg cac agc cag gca cgg ctg gcc gac cgc 432
 Leu Thr Val Glu Asp Trp Val His Ser Gln Ala Arg Leu Ala Asp Arg
 130 135 140

ctg ggc atc cag cag tgg gca gct atc gtc ggc ggt agc ctg ggt ggc 480
 Leu Gly Ile Gln Gln Trp Ala Ala Ile Val Gly Gly Ser Leu Gly Gly
 145 150 155 160

atg cag gcg ctg caa tgg acg atg acc tac ccc gag cgc gta cgc cac 528
 Met Gln Ala Leu Gln Trp Thr Met Thr Tyr Pro Glu Arg Val Arg His
 165 170 175

tgc gtc gac att gcc tcg gcc ccc aag ctg tcg gcg cag aac atc gcc 576
 Cys Val Asp Ile Ala Ser Ala Pro Lys Leu Ser Ala Gln Asn Ile Ala
 180 185 190

ttc aac gag gtg gcg cgt cag gcc att ctt acc gac cct gag tac cgc 624
 Phe Asn Glu Val Ala Arg Gln Ala Ile Leu Thr Asp Pro Glu Tyr Arg
 195 200 205

aga ggc tcg ttt cca gga cca ggt gtg atc ccc aag cgc ggc ctg atg 672
 Arg Gly Ser Phe Pro Gly Pro Gly Val Ile Pro Lys Arg Gly Leu Met
 210 215 220

ctg gca cgg atg gtc ggc cac att acc tat ctg tcc gat gat tcg atg 720
 Leu Ala Arg Met Val Gly His Ile Thr Tyr Leu Ser Asp Asp Ser Met
 225 230 235 240

ggt gaa aaa ttc ggc cga gag ctg aaa gcg aca agc tca act acg act 768
 Gly Glu Lys Phe Gly Arg Glu Leu Lys Ala Thr Ser Ser Thr Thr Thr
 245 250 255

tcc aca gcg tcg agt tcc agg tcg aaa gct acc tgc gct atc agg gcg 816
 Ser Thr Ala Ser Ser Ser Arg Ser Lys Ala Thr Cys Ala Ile Arg Ala
 260 265 270

agg agt ttt ccg gcc gtt tcg acg cca aca cct acc ttg atg acc aag 864
 Arg Ser Phe Pro Ala Val Ser Thr Pro Thr Pro Thr Leu Met Thr Lys
 275 280 285

gca ctg gac tat ttc gac ccg gcc gcc acg cac ggt ggt gat ctg gcc 912
 Ala Leu Asp Tyr Phe Asp Pro Ala Ala Thr His Gly Gly Asp Leu Ala
 290 295 300

gcc acc ctg gcc cac gtc acg gcg gac tac tgc atc tgt cgt tca cca 960
 Ala Thr Leu Ala His Val Thr Ala Asp Tyr Cys Ile Cys Arg Ser Pro
 305 310 315 320

ccg act gcg ctt ctc tcc ggc ccg ttc gcg cga gat cgt cga cgc gct 1008
 Pro Thr Ala Leu Leu Ser Gly Pro Phe Ala Arg Asp Arg Arg Arg Ala
 325 330 335

gat ggc cgc gcg caa gaa cgt ctg cta cct gga gat cga ttc gcc cta 1056
 Asp Gly Arg Ala Gln Glu Arg Leu Leu Pro Gly Asp Arg Phe Ala Leu
 340 345 350

cgg gca cga tgc att tcc tga 1077
 Arg Ala Arg Cys Ile Ser
 355

<210> 46

<211> 358

<212> PRT

<213> Pseudomonas putida

<400> 46

Met Ser Thr Val Phe Pro Glu Asp Ser Val Gly Leu Val Val Arg Gln
 1 5 10 15

Thr Ser Arg Phe Asp Glu Pro Leu Ala Leu Ala Cys Gly Arg Ser Leu

20						25						30					
Ala	Ser	Tyr	Glu	Leu	Val	Tyr	Glu	Thr	Tyr	Gly	Thr	Leu	Asn	Ala	Ser		
35						40						45					
Ala	Ser	Asn	Ala	Val	Leu	Ile	Cys	His	Ala	Leu	Ser	Gly	His	His	His		
50						55						60					
Ala	Ala	Gly	Tyr	His	Ala	Ala	Thr	Asp	Arg	Lys	Pro	Gly	Trp	Trp	Asp		
65			70						75			80					
Ser	Cys	Ile	Gly	Pro	Gly	Lys	Pro	Ile	Asp	Thr	Asn	Arg	Phe	Phe	Val		
			85						90			95					
Val	Ser	Leu	Asn	Asn	Leu	Gly	Gly	Cys	Asn	Gly	Ser	Thr	Gly	Pro	Ser		
			100			105						110					
Ser	Val	Asn	Pro	Ala	Thr	Gly	Lys	Pro	Tyr	Gly	Ala	Glu	Phe	Pro	Val		
115						120						125					
Leu	Thr	Val	Glu	Asp	Trp	Val	His	Ser	Gln	Ala	Arg	Leu	Ala	Asp	Arg		
130						135			140								
Leu	Gly	Ile	Gln	Gln	Trp	Ala	Ala	Ile	Val	Gly	Gly	Ser	Leu	Gly	Gly		
145			150						155			160					
Met	Gln	Ala	Leu	Gln	Trp	Thr	Met	Thr	Tyr	Pro	Glu	Arg	Val	Arg	His		
			165						170			175					
Cys	Val	Asp	Ile	Ala	Ser	Ala	Pro	Lys	Leu	Ser	Ala	Gln	Asn	Ile	Ala		
			180			185						190					
Phe	Asn	Glu	Val	Ala	Arg	Gln	Ala	Ile	Leu	Thr	Asp	Pro	Glu	Tyr	Arg		
195						200						205					
Arg	Gly	Ser	Phe	Pro	Gly	Pro	Gly	Val	Ile	Pro	Lys	Arg	Gly	Leu	Met		
210						215			220								
Leu	Ala	Arg	Met	Val	Gly	His	Ile	Thr	Tyr	Leu	Ser	Asp	Asp	Ser	Met		
225			230						235			240					
Gly	Glu	Lys	Phe	Gly	Arg	Glu	Leu	Lys	Ala	Thr	Ser	Ser	Thr	Thr	Thr		
			245						250			255					
Ser	Thr	Ala	Ser	Ser	Ser	Arg	Ser	Lys	Ala	Thr	Cys	Ala	Ile	Arg	Ala		
			260			265						270					
Arg	Ser	Phe	Pro	Ala	Val	Ser	Thr	Pro	Thr	Pro	Thr	Leu	Met	Thr	Lys		
275						280						285					
Ala	Leu	Asp	Tyr	Phe	Asp	Pro	Ala	Ala	Thr	His	Gly	Gly	Asp	Leu	Ala		
290						295						300					
Ala	Thr	Leu	Ala	His	Val	Thr	Ala	Asp	Tyr	Cys	Ile	Cys	Arg	Ser	Pro		
305			310						315			320					
Pro	Thr	Ala	Leu	Leu	Ser	Gly	Pro	Phe	Ala	Arg	Asp	Arg	Arg	Arg	Ala		
			325						330			335					
Asp	Gly	Arg	Ala	Gln	Glu	Arg	Leu	Leu	Pro	Gly	Asp	Arg	Phe	Ala	Leu		
			340			345						350					
Arg	Ala	Arg	Cys	Ile	Ser												

355

<210> 47

<211> 52

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 47

cccgggatcc gctagcggcg cgccggccgg cccgggtgtga aataccgcac ag 52

<210> 48

<211> 53

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 48

tctagactcg agcggcccgcg gccggccttt aaattgaaga cgaaagggcc tcg 53

<210> 49

<211> 47

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 49

gagatctaga cccggggatc cgctagcggg ctgctaaagg aagcgga 47

<210> 50

<211> 38

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 50

gagaggcgcg ccgctagcgt gggcgaagaa ctccagca 38

<210> 51

<211> 34

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 51

gagagggcgcg ccgcgcaaag tcccgttcg tgaa 34

<210> 52
<211> 34
<212> DNA
<213> Künstliche Sequenz

<220>
<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 52
gagagggcg cgcgtcaagt cggcgaagcc acgc

34

<210> 53
<211> 140
<212> DNA
<213> Künstliche Sequenz

<220>
<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 53
tcgaatttaa atctcgagag gcctgacgctc gggcccggtta ccacgcgtca tatgactagt 60
tcggacctag ggatatcgctc gacatcgatg ctcttctgcg ttaattaaca attgggatcc 120
tctagaccgc ggattttaa 140

<210> 54
<211> 140
<212> DNA
<213> Künstliche Sequenz

<220>
<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 54
gatcatttaa atccccgggtc tagaggatcc caattgttaa ttaacgcaga agagcatcga 60
tgctgacgat atccctaggt ccgaactagt catatgacgc gtggtaccgc gcccgacgctc 120
aggcctctcg agattttaa 140

<210> 55
<211> 33
<212> DNA
<213> Künstliche Sequenz

<220>
<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 55
gagagcggcc gccgacctt ttaacccat cac

33

<210> 56
<211> 32
<212> DNA
<213> Künstliche Sequenz

<220>
<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 56
aggagcggcc gccatcggca ttttcttttg cg

32

<210> 57

<211> 5091

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz: Plasmid

<400> 57

```
gcccgcactg ccttcgcgaa gccttgcccc gcggaaattt cctccaccga gttcgtgcac 60
acccctatgc caagcttctt tcaccctaaa ttcgagagat tggattctta ccgtggaaat 120
tcttcgcaaa aatcgcccc tgatcgccct tgcgacgttg gcgtcggtgc cgctggttgc 180
gcttggtctg accgacttga tcagcggccg ctcgatttaa atctcgagag gcctgacgtc 240
gggcccggta ccacgcgtca tatgactagt tcggacctag ggatatcgtc gacatcgatg 300
ctcttctgcy ttaattaaca attgggatcc tctagacctg ggatttaa at cgctagcggg 360
ctgctaaagg aagcggaaaca cgtagaaagc cagtccgcag aaacggtgct gaccccgat 420
gaatgtcagc tactgggcta tctggacaag ggaaaacgca agcgcaaaga gaaagcaggt 480
agcttgcaat gggcttacat ggcgatagct agactgggcy gttttatgga cagcaagcga 540
accggaattg ccagctgggg cgccctctgg taagggtggg aagccctgca aagtaaaactg 600
gatggctttc ttgccgcaa ggatctgatg gcgcagggga tcaagatctg atcaagagac 660
aggatgagga tgcgttcgca tgattgaaca agatggattg cagcagggtt ctccggccgc 720
ttgggtggag aggctattcg gctatgactg ggcacaacag acaatcggtc gctctgatgc 780
cgccgtgttc cggctgtcag cgcaggggcg cccggttctt tttgtcaaga ccgacctgtc 840
cgggtgccctg aatgaactgc aggacgaggc agcgcggcta tcgtggctgg ccacgacggg 900
cgttcccttg gcagctgtgc tcgacgttgt cactgaagcg ggaagggact ggctgctatt 960
gggcgaagtg ccggggcagg atctcctgtc atctcacctt gtcctgccc agaaagtatc 1020
catcatggct gatgcaatgc ggcggctgca tacgcttgat ccggctacct gccattcga 1080
ccaccaagcg aaacatcgca tcgagcgagc acgtactcgg atggaagccg gtcttgtcga 1140
tcaggatgat ctggacgaag agcatcaggg cgtcgcgcca gccgaactgt tcgccaggct 1200
caaggcgcgc atgccgcagc gcgaggatct gcctcgctgc catggcgatg cctgcttgcc 1260
gaatatcatg gtggaaaatg gccgcttttc tggattcatc gactgtggcc ggctgggtgt 1320
ggcggaccgc tatcaggaca tagcgttggc taccctgat attgctgaag agcttggcgg 1380
cgaatgggct gaccgcttcc tcgtgcttta cggatcgcc gctcccgat cgcagcgcac 1440
cgcttcttat cgcttcttg acgagttctt ctgagcggga ctctggggtt cgaaatgacc 1500
gaccaagcga cgcccaacct gccatcacga gatttcgatt ccaccgccgc cttctatgaa 1560
aggttgggct tcggaatcgt ttccgggac gcccgctgga tgatcctcca gcgcggggat 1620
agctcttg agttcttcgc ccacgttagc ggcgcgcgg ccggcccgggt gtgaaatacc 1680
gcacagatgc gtaaggagaa aataccgat caggcgctct tccgcttctc cgctcactga 1740
ctcgctgcgc tcggtcgttc ggctgcggcg agcggtatca gctcactcaa aggcggtaat 1800
acggttatcc acagaatcag gggataacgc aggaaagaac atgtgagcaa aaggccagca 1860
aaaggccagg aaccgtaaaa aggcgcggtt gctggcggtt ttccataggc tccgcccccc 1920
tgacgagcat caaaaaatc gacgtcaag tcagaggtgg cgaaaccga caggactata 1980
aagataccag gcgtttcccc ctgttctccc ttccgggaagc gtggcgcttt ctcatagctc 2100
gcttaccgga tacctgtccg cgggtgtaggt cgctcgctcc aagctgggct gtgtgcacga 2160
acccccggt cagccgacc gctgcgcctt atccggtaac tatcgtcttg agtccaaccc 2220
ggtaagacac gacttatcgc cactggcagc agccactggt aacaggatta gcagagcgag 2280
gtatgtaggc ggtgctacag agttcttgaa gtggtggcct aactacggct acactagaag 2340
gacagtattt ggtatctgcg ctctgctgaa gccagttacc ttccgaaaaa gaggtagtag 2400
ctcttgatcc ggcaaacaaa ccaccgtgg tagcgggtgt ttttttgttt gcaagcagca 2460
gattacgcgc agaaaaaaag gatctcaaga agatcctttg atcttttcta cggggtctga 2520
cgctcagtgg aacgaaaact caggttaagg gattttggtc atgagattat caaaaaggat 2580
cttcacctag atccttttaa aggcgggccc cgcccgcgca aagtcgccgt tcgtgaaaaa 2640
ttctgtgccc cgtgatttcc cgccaaaaac tttaacgaac gttcgttata atggtgtcat 2700
gaccttcacg acgaagtact aaaattggcc cgaatcatca gctatggatc tctctgatgt 2760
cgcgctggag tccgacgcgc tcgatgctgc atcacgagac tgggccagtg ccgcgagcga 2880
ccgagctctc gatacgacgg acgcgccagc gctggctgac gagctgcgtg ctccggccagc 2940
cctagaaact ctctggcgcg atcttgagga aatcagttgc gcctactgcg gtggcctgat 3000
gccaggagga cgcacagtag tggaggatgc cgcgaaatat tgctcagatg cgtgtcgtgc 3060
tcctccccgg cctgaccgcg ccaacaaacg ccacgcccag gagctggagg cggctagggtc 3120
cgcaatggcg ctggaagtgc gtcccccgag cgaaattttg gccatggtcg tcacagagct 3180
```

```

ggaagcggca gcgagaatta tcgcgatcgt ggcgggtgcc gcaggcatga caaacatcgt 3240
aaatgccgcg tttcgtgtgc cgtggcgcgc caggacgtgt cagcgccgcc accacctgca 3300
ccgaatcggc agcagcgtcg cgcgtcgaaa aagcgcacag gcggcaagaa gcgataagct 3360
gcacgaatac ctgaaaaatg ttgaacgccc cgtgagcggg aactcacagg gcgtcggcta 3420
acccccagtc caaacctggg agaaagcgct caaaaatgac tctagcggat tcacgagaca 3480
ttgacacacc ggcttggaag ttttccgctg atctgttcga caccatccc gagctcgcgc 3540
tgcatcacg tggctggacg agcgaagacc gccgcgaatt cctcgctcac ctgggcagag 3600
aaaatttcca gggcagcaag acccgcgact tcgccagcgc ttggatcaaa gaccgggaca 3660
cggagaaca cagccgaagt tataccgagt tggttcaaaa tcgcttgccc ggtgccagta 3720
tgttgctctg acgcacgcgc agcacgcagc cgtgcttgte ctggacattg atgtgccgag 3780
ccaccaggcc ggcgggaaaa tcgagcacgt aaaccccgag gtctacgcga ttttggagcg 3840
ctgggcacgc ctggaaaaag cgccagcttg gatcggcgtg aatccactga gcgggaaatg 3900
ccagctcatc tggctcattg atccgggtgta tgccgcagca ggcatgagca gccgaatat 3960
gcgcctgctg gctgcaacga ccgaggaaat gaccgcgctt ttcggcgtg accaggcttt 4020
ttcacatagg ctgagccgtg gccactgcac tctccgacga tcccagcgt accgctggca 4080
tgcccagcac aatcgcgctg atcgccctagc tgatcttatg gaggttgctc gcatgatctc 4140
aggcacagaa aaacctaaaa aacgctatga gcaggagttt tctagcggac gggcacgtat 4200
cgaagcggca agaaaagcca ctgcggaagc aaaagcactt gccacgcttg aagcaagcct 4260
gccgagcggc gctgaagcgt ctggagagct gatcgacggc gtccgtgtcc tctggactgc 4320
tccaggcgct gccgcccgtg atgagacggc ttttcgccac gctttgactg tgggatacca 4380
gttaaaagcg gctggtgagc gcctaaaaga caccaagggt catcgagcct acgagcgtgc 4440
ctacaccgtc gctcaggcgg tcggaggagg ccgtgagcct gatctgccgc cggactgtga 4500
ccgccagacg gattggccgc gacgtgtgcg cggctacgtc gctaaaggcc agccagtcgt 4560
ccctgctcgt cagacagaga cgagagcca gccgaggcga aaagctctgg ccactatggg 4620
aagacgtggc ggtaaaaagg ccgcagaacg ctggaaagac ccaaacagtg agtacgcccg 4680
agcacagcga gaaaaactag ctaagtccag tgttgaggga gagactggct cgtggccgac 4800
tcgcttgacc attgcagggt ggtttatgac gtcacgtcag accgtgaata gagcacttaa 4860
aatcaatgaa gctatgtctg aatttagcgt gccgaagct tcccagtaaa tgtgccatct 4920
ggtctgcggg cattgaactt ccacgaggac ctctctcttg gcctcctttc taggtcgggc 4980
cgtaggcaga aaacggttcc cccgtagggt ctctctcttg gcctcctttc taggtcgggc 4980
tgattgctct tgaagctctc tagggggggt cacaccatag gcagataacg ttccccaccg 5040
gctcgcctcg taagcgcaca aggactgctc ccaaagatct tcaaagccac t 5091

```

<210> 58

<211> 4323

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:Plasmid

<400> 58

```

tctctcagcg tatggttgct gcctgagctg tagttgcctt catcgatgaa ctgctgtaca 60
ttttgatacg tttttccgtc accgtcaaag attgatttat aatcctctac accgttgatg 120
ttcaaagagc tgtctgatgc tgatacgtaa acttggtgcag ttgtcagtg ttgtttgccg 180
taatgtttac cggagaaatc agtgtagaat aaacggattt ttcgctcaga tgtaaatgtg 240
gctgaacctg accattcttg tgtttggtct tttaggatag aatcatttgc atcgaatttg 300
tcgctgtctt taaagacgcg gccagcgttt tccagctgt caatagaagt ttcgccgact 360
ttttgataga acatgtaaat cgatgtgtca tccgcatttt taggatctcc ggctaatagca 420
aagacgatgt ggtagccgtg atagtttgcy acagtgcctg cagcgttttg taatggccag 480
ctgtcccaaa cgtccaggcc ttttgacaaa gagatatttt taattgtgga cgaatcaaat 540
tcagaaactt gatatttttc atttttttgc tgttcaggga tttgcagcat atcatggcgt 600
gtaatatggg aaatgccgta tgtttcctta tatggctttt ggttcgtttc tttcgcaaac 660
gcttgagttg cgctcctctg cagcagtcgc gtatgaaagg ttagtactgt tgcttgtttt 720
gcaaactttt tgatgttcat cgttcatgtc tcctttttta tgtactgtgt tagcggctctg 780
cttcttccag cctcctgttt tgaagatggc aagttagtta cgcacaataa aaaaagacct 840
aaaatatgta aggggtgacg ccaaagtata cactttgccc tttacacatt ttaggtcttg 900
cctgctttat cagtaacaaa ccgcgcgat ttacttttcg acctattct attagactct 960
cgtttgatt gcaactggtc tattttcttc ttttgtttga tagaaaatca taaaaggatt 1020
tgcagactac gggcctaaag aactaaaaaa tctatctgtt tcttttcatt ctctgtattt 1080
tttatagttt ctgttgcatg ggcataaagt tgccttttta atcacaattc agaaaatata 1140

```



```

ataatatctc atttcactaa ataatagtga acggcaggta tatgtgatgg gttaaaaagg 1200
atcggcgggc gctcgattta aatctcgaga ggcctgacgt cgggcccggg accacgcgtc 1260
atatgactag ttcggaccta gggatatcgt cgacatcgat gctcttctgc gttaattaac 1320
aattgggata ctctagaccc gggattttaa tcgctagcgg gctgctaaag gaagcggaac 1380
acgtagaaaag ccagtcgcga gaaacgggtg tgaccccgga tgaatgtcag ctactgggct 1440
atctggacaa gggaaaacgc aagcgcaaag agaaagcagg tagcttgtag tgggcttaca 1500
tggcgatagc tagactgggc ggttttatgg acagcaagcg aaccggaatt gccagctggg 1560
gcgcctctcg gtaagggttg gaagccctgc aaagtaaact ggatggcttt cttgcgcgca 1620
aggatctgat ggcgcagggt atcaagatct gatcaagaga caggatgagg atcgtttcgc 1680
atgattgaac aagatggatt gcacgcagg tctccggccg cttgggtgga gaggctattc 1740
ggctatgact gggcacaaac gacaatcggc tgctctgtag ccgctgtgtt ccggctgtca 1800
gcgcaggggc gcccggttct ttttgtcaag accgacctgt ccggtgccct gaatgaactg 1860
caggacgagg cagcgcggct atcgtggctg gccacgacgg cgttcccttg cgcagctgtg 1920
ctcgacgttg tcaactgaag gggaaaggac tggctgctat tgggcgaagt gccggggcag 1980
gatctcctgt catctcacct tgctcctgcc gagaaagtat ccatcatggc tgatgcaatg 2040
cggcggtcgc atacgcttga tccggctacc tgcccattcg accaccaagc gaaacatcgc 2100
atcgagcgag cacgtactcg gatggaagcc ggtcttgtcg atcaggatga tctggacgaa 2160
gagcatcagg ggtcgcgcgc agccgaactg ttcgccaggc tcaaggcgcg catgcccgac 2220
ggcgaggatc tcgtcgtgac ccatggcgat gcctgcttgc cgaatatcat ggtggaaaat 2280
ggcgctttt cttgattcat cgactgtggc cggctgggtg tggcggaccg ctatcaggac 2340
atagcggttg ctaccctgta tattgtgtaa gagcttggcg gcgaatgggc tgaccgcttc 2400
ctcgtgcttt acggtatcgc cgtcccgat tcgcagcgca tcgcttctta tcgcttctt 2460
gacgagttct tctgagcggg actctggggg tcgaaatgac cgaccaagcg acgcccacc 2520
tgccatcacg agatttcgat tccaccgcgc ccttctatga aaggttgggc ttcggaatcg 2580
ttttccggga cgccggctgg atgatcctcc agcgcgggga tctcatgctg gagttcttcg 2640
cccacgctag cggcgcgccg gccggccccg tgtgaaatac cgcacagatg cgtaggaga 2700
aaataccgca tcaggcgctc ttccgcttcc tcgctcactg actcgctgcg ctcggtcgtt 2760
cggctgcggc gagcggtatc agctcactca aaggcggtaa tacggttatc cacagaatac 2820
ggggataaac caggaaagaa catgtgagca aaaggccagc gaaccgtaaa 2880
aaggccgcgt tgctggcggt tttccatagg ctccgcccc ctgacgagca tcacaaaaat 2940
cgacgctcaa gtcagaggtg gcgaaacccg acaggactat aaagatacca ggcgtttccc 3000
cctggaagct ccctcgtgcg ctctcctgtt ccgaccctgc cgcttaccgg atacctgtcc 3060
gcctttctcc cttcggaag cgtggcgctt tctcatagct cacgctgtag gtatctcagt 3120
tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg aacccccgt tcagcccgac 3180
cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc cggtagaca cgacttatcg 3240
ccactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg cgggtctaca 3300
gagttcttga agtggtggcc taactacggc tacactagaa ggacagtatt tggatatcgc 3360
gctctgctga agccagttac cttcggaaaa agagtgggta gctcttgatc cggcaaaaa 3420
accaccgctg gtagcgggtg tttttttgtt tgcaagcagc agattacgcg cagaaaaaaa 3480
ggatctcaag aagatccttt gatcttttct acggggtctg acgctcagtg gaacgaaaa 3540
tcacgttaag ggatttttgt catgagatta tcaaaaagga tcttcaccta gatcctttta 3600
aaggccggcc gcggccgcca tcggcatttt cttttgcgtt tttatttgtt aactgttaat 3660
tgtccttgtt caaggatgct gtctttgaca acagatgttt tcttgccctt gatgttcagc 3720
aggaagctcg gcgcaaactg tgattgtttg tctgcgtaga atcctctgtt tgtcatatag 3780
cttgtaatca cgactattgt tcctttcgct tgaggtacag cgaagtgtga gtaagtaaag 3840
gttacatcgt taggatcaag atccattttt aacacaaggc cagttttgtt cagcggcttg 3900
tatgggccag ttaaagaatt agaaacataa ccaagcatgt aaatatcgtt agacgtaatg 3960
ccgtcaatcg tcattttttg tccgcgggag tcagtgaaca ggtaccattt gccgttcatt 4020
ttaaagacgt tcgcgcgttc aatttcactt gttactgtgt tagatgcaat cagcggtttc 4080
atcacttttt tcagtgtgta atcatcgctt agctcaatca taccgagagc gccgtttgct 4140
aactcagccg tgcgtttttt atcgctttgc agaagttttt gactttcttg acggaagaat 4200
gatgtgcttt tgccatagta tgctttgtta aataaagatt cttcgccttg gtagccatct 4260
tcagttccag tgtttgcttc aaataactaag tatttgtggc ctttatcttc tacgtagtga 4320
gga

```

<210> 59

<211> 35

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:PCR Primer

<400> 59
gagagagaga cgcgtcccag tggctgagac gcatc 35

<210> 60
<211> 34
<212> DNA
<213> Künstliche Sequenz

<220>
<223> Beschreibung der künstlichen Sequenz:PCR Primer

<400> 60
ctctctctgt cgacgaattc aatcttacgg cctg 34

<210> 61
<211> 5860
<212> DNA
<213> Künstliche Sequenz

<220>
<223> Beschreibung der künstlichen Sequenz:Plasmid

<400> 61
cccgggtacca cgcgtcccag tggctgagac gcatccgcta aagccccagg aaccctgtgc 60
agaaagaaaa cactcctctg gctaggtaga cacagtttat aaaggtagag ttgagcgggt 120
aactgtcagc acgtagatcg aaagggtgcac aaagggtggcc ctggctcgta agaaatatgg 180
cgggttcctcg cttgagagtg cggaacgcat tagaaacgtc gctgaacgga tcgttgccac 240
caagaaggct ggaaatgatg tcgtggttgt ctgctccgca atgggagaca ccacggatga 300
acttctagaa cttgcagcgg cagtgaatcc cgttccgcca gctcgtgaaa tggatatgct 360
cctgactgct ggtgagcgta tttctaacgc tctcgtcgcc atggctattg agtcccttgg 420
cgcagaagcc caatctttca cgggctctca ggctgggtgtg ctcaccaccg agcgccacgg 480
aaacgcacgc attgttgatg tcaactccagg tcgtgtgctg gaagcactcg atgagggcaa 540
gatctgcatt gttgctggtt tccagggtgt taataaagaa acccgcatg tcaccacgtt 600
gggtcgtggt ggttctgaca ccaactgcagt tgcgttggca gctgctttga acgctgatgt 660
gtgtgagatt tactcggacg ttgacggtgt gtataccgct gacccgcgca tcgttcctaa 720
tgcacagaag ctggaaaagc tcagcttcga agaaatgctg gaacttgctg ctgttggttc 780
caagattttg gtgctgcgca gtgttgaata cgctcgtgca ttcaatgtgc cacttcgcgt 840
acgctcgtct tatagtaatg atcccggcac tttgattgcc ggctctatgg aggatattcc 900
tgtggaagaa gcagtcctta ccgggtgtcg aaccgacaag tccgaagcca aagtaaccgt 960
tctgggtatt tccgataagc caggcgaggc tgcgaagggt ttccgtgctg tggctgatgc 1020
agaaatcaac attgacatgg ttctgcagaa cgtctcttct gtagaagacg gcaccaccga 1080
catcaccttc acctgcctc gttccgacgg ccgccgcgcg atggagatct tgaagaagct 1140

tcagggttcag ggcaactgga ccaatgtgct ttacgacgac caggtcggca aagtctccct 1200
cgtgggtgct ggcataaagt ctacccagg tgttaccgca gagttcatgg aagctctgcg 1260
cgatgtcaac gtgaacatcg aattgatttc cacctctgag attcgtatatt ccgtgctgat 1320
ccgtgaagat gatctggatg ctgctgcacg tgcattgcat gagcagttcc agctgggagg 1380
cgaagacgaa gccgtcgttt atgcaggcac cggacgctaa agttttaaag gagtagtttt 1440
acaatgacca ccatcgcagt tgttggtgca accggccagg tcggccagggt tatgcgaccc 1500
cttttggaag agcgcaattt ccagctgac actgttcggt tctttgcttc cccacgttcc 1560
gcaggccgta agattgaatt cgtcgacatc gatgctcttc tgcgttaatt aacaattggg 1620
atcctctaga cccgggattt aaatcgctag cgggctgcta aaggaagcgg aacacgtaga 1680
aagccagtcc gcagaaacgg tgctgacccc ggatgaatgt cagctactgg gctatctgga 1740
caagggaaaa cgcaagcgca aagagaaagc aggtagcttg cagtgggctt acatggcgat 1800
agctagactg ggcggtttta tggacagcaa gcgaaccgga attgccagct ggggcgcctt 1860
ctggtaaggt tgggaagccc tgcaaagtaa actggatggc tttcttgccg ccaaggatct 1920
gatggcgagc gggatcaaga tctgatcaag agacaggatg aggatcgttt cgcattgattg 1980
aacaagatgg attgcacgca gggtctccgg ccgcttgggt ggagaggcta ttcggctatg 2040
actgggcaca acagacaatc ggctgctctg atgccgccgt gttccggctg tcagcgagc 2100
ggcgcccggt tctttttgtc aagaccgacc tgtccgggtg cctgaatgaa ctgcaggacg 2160
aggcagcgcg gctatcgtgg ctggccacga cgggcgttcc ttgcgagct gtgctcgacg 2220
ttgtcactga agcgggaagg gactggctgc tattgggcga agtgccgggg caggatctcc 2280
tgtcatctca ccttgctcct gccgagaaag tatccatcat ggctgatgca atgcggcggc 2340
tgcatacgct tgatccggct acctgcccat tcgaccacca agcgaacat cgcacgagc 2400
gagcacgtac tcggatggaa gccggtcttg tcgatcagga tgatctggac gaagagcatc 2460
aggggctcgc gccagccgaa ctgttcgcca ggctcaaggc gcgcatgcc cagggcgagg 2520
atctcgctgt gacccatggc gatgcctgct tgccgaatat catgggtggaa aatggccgct 2580
tttctggatt catcgactgt ggccggctgg gtgtggcgga ccgctatcag gacatagcgt 2640
tggctacccg tgatattgct gaagagcttg gcggcgaaat ggctgaccgc ttctctgtgc 2700
tttacggtat cgcgctccc gattcgcagc gcatcgctt ctatcgctt cttgacgagt 2760
tcttctgagc gggactctgg ggttcgaaat gaccgaccaa gcgacgcca acctgccatc 2820
acgagatttc gattccaccg ccgccttcta tgaaagggtg ggcttcggaa tcgttttccg 2880
ggacgccggc tggatgatcc tccagcgcg ggatctcatg ctggagttct tcgcccacgc 2940
tagcggcgcg ccggccggcc cgggtgtgaaa taccgcacag atgcgtaagg agaaaatacc 3000

gcacacaggcg ctcttccgct tectcgctca ctgactcgct gcgctcggtc gttcgggtgc 3060
ggcgagcggt atcagctcac tcaaaggcgg taatacgggt atccacagaa tcaggggata 3120
acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg 3180
cgttgctggc gtttttccat aggctccgcc cccctgacga gcacacaaa aatcgacgct 3240
caagtcagag gtggcgaaac ccgacaggac tataaagata ccaggcgttt cccctggaa 3300
gctccctcgt gcgctctcct gttccgacct tgcgcttac cggatacctg tccgccttcc 3360
tcccttcggg aagcgtggcg ctttctcata gctcacgctg taggtatctc agttcgggtg 3420
aggctcgttc ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg 3480
ccttatccgg taactatcgt cttgagtcca acccggttaag acacgactta tcgccactgg 3540
cagcagccac tggtaacagg attagcagag cgaggatatgt aggcgggtgt acagagttct 3600
tgaagtgggt gcctaactac ggctacacta gaaggacagt atttggtatc tgcgctctgc 3660
tgaagccagt taccttcgga aaaagagttg gtagctcttg atccggcaaa caaacaccg 3720
ctggtagcgg tggttttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc 3780
aagaagatcc tttgatcttt tctacggggt ctgacgctca gtggaacgaa aactcacggt 3840
aagggtttt ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt ttaaaggccg 3900
gccgcggccg ccacgcgcat tttcttttgc gtttttatct gttaactgtt aattgtcctt 3960
gttcaaggat gctgtctttg acaacagatg ttttcttgcc tttgatgttc agcaggaagc 4020
tcggcgcaaa cgttgattgt ttgtctgcgt agaatcctct gtttgtcata tagcttgtaa 4080
tcacgacatt gtttcccttc gcttgaggta cagcgaagtg tgagtaagta aaggttacat 4140
cgttaggatc aagatccatt tttaacacaa ggccagtttt gttcagcggc ttgtatgggc 4200
cagttaaaga attagaaaca taaccaagca tgtaaatac gttagacgta atgccgtcaa 4260
tcgtcatctt tgatccgcgg gagtcagtga acaggtacca tttgccgttc attttaaaga 4320
cgttcgcgcg ttcaatttca tctgttactg tgttagatgc aatcagcggg ttcacactt 4380
ttttcagtgt gtaatcatcg tttagctcaa tcataccgag agcgccggtt gctaactcag 4440
ccgtgcgttt tttatcgctt tgcagaagtt tttgacttcc ttgacggaag aatgatgtgc 4500
ttttgccata gtatgctttg tttaataaag attcttcgcc ttggtagcca tcttcagtcc 4560
cagtgtttgc ttcaaatact aagtatttgt ggcttttacc ttctacgtag tgaggatctc 4620
tcagcgatat gttgtcgctt gagctgtagt tgccttcacc gatgaactgc tgtacatttt 4680
gatacgtttt tccgtcaccg tcaaagattg atttataatc ctctacaccg ttgatgttca 4740
aagagctgtc tgatgctgat acgttaactt gtgcagttgt cagtgtttgt ttgccgtaat 4800
gtttaccgga gaaatcagtg tagaataaac ggatttttcc gtcagatgta aatgtggctg 4860
aacctgacca ttcttgtgtt tggctcttta ggatagaatc atttgcacg aatttgtcgc 4920

```

tgtcttttaaa gacgcggcca gcgtttttcc agctgtcaat agaagtttcg ccgacttttt 4980
gatagaacat gtaaategat gtgtcatccg catttttagg atctccggct aatgcaaaga 5040
cgatgtggta gccgtgatag tttgcgacag tgccgtcagc gttttgtaat ggccagctgt 5100
cccaaacgtc caggcctttt gcagaagaga tatttttaat tgtggacgaa tcaaattcag 5160
aaacttgata tttttcattt ttttgctggt cagggatttg cagcatatca tggcgtgtaa 5220
tatgggaaat gccgtatggt tccttatatg gcttttggtt cgtttctttc gcaaacgctt 5280
gagttgcgcc tcctgccagc agtgcggtag taaagggtta tactgttgct tgttttgcaa 5340
actttttgat gttcatcggt catgtctcct tttttatgta ctgtgttagc ggtctgcttc 5400
ttccagccct cctgtttgaa gatggcaagt tagttacgca caataaaaaa agacctaaaa 5460
tatgtaaggg gtgacgcca agtatacact ttgcccttta cacatttttag gtcttgcttg 5520
ctttatcagt aacaaacccg cgcgatttac ttttcgacct cattctatta gactctcggt 5580
tggattgcaa ctggtctatt ttcctctttt gtttgataga aaatcataaa aggatttgca 5640
gactacgggc ctaaagaact aaaaaatcta tctgtttctt ttcattctct gtatttttta 5700
tagtttctgt tgcattgggca taaagttgcc tttttaatca caattcagaa aatatcataa 5760
tatctcattt cactaaataa tagtgaacgg caggtatatg tgatgggtta aaaaggatcg 5820
gcggccgctc gatttaaate tcgagaggcc tgacgtcggg 5860

```

<210> 62

<211> 38

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:PCR Primer

<400> 62

cggcaccacc gacatcatct tcacctgccc tcgttccg

38

<210> 63

<211> 38

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:PCR Primer

<400> 63

cggaacgagg gcagggtgaag atgatgtcgg tgggtgccg

38

<210> 64

<211> 1266

<212> DNA

<213> LysC Mutante

<220>

<221> CDS

<222> (1) .. (1266)

<223>

<400> 64

gtg gcc ctg gtc gta cag aaa tat ggc ggt tcc tcg ctt gag agt gcg	48
Val Ala Leu Val Val Gln Lys Tyr Gly Gly Ser Ser Leu Glu Ser Ala	
1 5 10 15	
gaa cgc att aga aac gtc gct gaa cgg atc gtt gcc acc aag aag gct	96
Glu Arg Ile Arg Asn Val Ala Glu Arg Ile Val Ala Thr Lys Lys Ala	
20 25 30	
gga aat gat gtc gtg gtt gtc tgc tcc gca atg gga gac acc acg gat	144
Gly Asn Asp Val Val Val Val Cys Ser Ala Met Gly Asp Thr Thr Asp	
35 40 45	
gaa ctt cta gaa ctt gca gcg gca gtg aat ccc gtt ccg cca gct cgt	192
Glu Leu Leu Glu Leu Ala Ala Val Asn Pro Val Pro Pro Ala Arg	
50 55 60	
gaa atg gat atg ctc ctg act gct ggt gag cgt att tct aac gct ctc	240
Glu Met Asp Met Leu Leu Thr Ala Gly Glu Arg Ile Ser Asn Ala Leu	
65 70 75 80	
gtc gcc atg gct att gag tcc ctt ggc gca gaa gcc caa tct ttc acg	288
Val Ala Met Ala Ile Glu Ser Leu Gly Ala Glu Ala Gln Ser Phe Thr	
85 90 95	
ggc tct cag gct ggt gtg ctc acc acc gag cgc cac gga aac gca cgc	336
Gly Ser Gln Ala Gly Val Leu Thr Thr Glu Arg His Gly Asn Ala Arg	
100 105 110	
att gtt gat gtc act cca ggt cgt gtg cgt gaa gca ctc gat gag ggc	384
Ile Val Asp Val Thr Pro Gly Arg Val Arg Glu Ala Leu Asp Glu Gly	
115 120 125	
aag atc tgc att gtt gct ggt ttc cag ggt gtt aat aaa gaa acc cgc	432
Lys Ile Cys Ile Val Ala Gly Phe Gln Gly Val Asn Lys Glu Thr Arg	
130 135 140	
gat gtc acc acg ttg ggt cgt ggt ggt tct gac acc act gca gtt gcg	480
Asp Val Thr Thr Leu Gly Arg Gly Gly Ser Asp Thr Thr Ala Val Ala	
145 150 155 160	
ttg gca gct gct ttg aac gct gat gtg tgt gag att tac tcg gac gtt	528
Leu Ala Ala Ala Leu Asn Ala Asp Val Cys Glu Ile Tyr Ser Asp Val	
165 170 175	
gac ggt gtg tat acc gct gac ccg cgc atc gtt cct aat gca cag aag	576
Asp Gly Val Tyr Thr Ala Asp Pro Arg Ile Val Pro Asn Ala Gln Lys	
180 185 190	
ctg gaa aag ctc agc ttc gaa gaa atg ctg gaa ctt gct gct gtt ggc	624
Leu Glu Lys Leu Ser Phe Glu Glu Met Leu Glu Leu Ala Ala Val Gly	
195 200 205	
tcc aag att ttg gtg ctg cgc agt gtt gaa tac gct cgt gca ttc aat	672
Ser Lys Ile Leu Val Leu Arg Ser Val Glu Tyr Ala Arg Ala Phe Asn	
210 215 220	
gtg cca ctt cgc gta cgc tcg tct tat agt aat gat ccc ggc act ttg	720
Val Pro Leu Arg Val Arg Ser Ser Tyr Ser Asn Asp Pro Gly Thr Leu	

225	230	235	240	
att gcc ggc tct atg gag gat att cct gtg gaa gaa gca gtc ctt acc				768
Ile Ala Gly Ser Met Glu Asp Ile Pro Val Glu Glu Ala Val Leu Thr				
	245	250	255	
ggt gtc gca acc gac aag tcc gaa gcc aaa gta acc gtt ctg ggt att				816
Gly Val Ala Thr Asp Lys Ser Glu Ala Lys Val Thr Val Leu Gly Ile				
	260	265	270	
tcc gat aag cca ggc gag gct gcg aag gtt ttc cgt gcg ttg gct gat				864
Ser Asp Lys Pro Gly Glu Ala Ala Lys Val Phe Arg Ala Leu Ala Asp				
	275	280	285	
gca gaa atc aac att gac atg gtt ctg cag aac gtc tct tct gta gaa				912
Ala Glu Ile Asn Ile Asp Met Val Leu Gln Asn Val Ser Ser Val Glu				
	290	295	300	
gac ggc acc acc gac atc atc ttc acc tgc cct cgt tcc gac ggc cgc				960
Asp Gly Thr Thr Asp Ile Ile Phe Thr Cys Pro Arg Ser Asp Gly Arg				
	305	310	315	320
cgc gcg atg gag atc ttg aag aag ctt cag gtt cag ggc aac tgg acc				1008
Arg Ala Met Glu Ile Leu Lys Lys Leu Gln Val Gln Gly Asn Trp Thr				
	325	330	335	
aat gtg ctt tac gac gac cag gtc ggc aaa gtc tcc ctc gtg ggt gct				1056
Asn Val Leu Tyr Asp Asp Gln Val Gly Lys Val Ser Leu Val Gly Ala				
	340	345	350	
ggc atg aag tct cac cca ggt gtt acc gca gag ttc atg gaa gct ctg				1104
Gly Met Lys Ser His Pro Gly Val Thr Ala Glu Phe Met Glu Ala Leu				
	355	360	365	
cgc gat gtc aac gtg aac atc gaa ttg att tcc acc tct gag att cgt				1152
Arg Asp Val Asn Val Asn Ile Glu Leu Ile Ser Thr Ser Glu Ile Arg				
	370	375	380	
att tcc gtg ctg atc cgt gaa gat gat ctg gat gct gct gca cgt gca				1200
Ile Ser Val Leu Ile Arg Glu Asp Asp Leu Asp Ala Ala Ala Arg Ala				
	385	390	395	400
ttg cat gag cag ttc cag ctg ggc ggc gaa gac gaa gcc gtc gtt tat				1248
Leu His Glu Gln Phe Gln Leu Gly Gly Glu Asp Glu Ala Val Val Tyr				
	405	410	415	
gca ggc acc gga cgc taa				1266
Ala Gly Thr Gly Arg				
	420			

<210> 65
 <211> 421
 <212> PRT
 <213> LysC Mutante

<400> 65

Val Ala Leu Val Val Gln Lys Tyr Gly Gly Ser Ser Leu Glu Ser Ala
1 5 10 15

Glu Arg Ile Arg Asn Val Ala Glu Arg Ile Val Ala Thr Lys Lys Ala

20

25

30

Gly Asn Asp Val Val Val Val Cys Ser Ala Met Gly Asp Thr Thr Asp
35 40 45

Glu Leu Leu Glu Leu Ala Ala Ala Val Asn Pro Val Pro Pro Ala Arg
50 55 60

Glu Met Asp Met Leu Leu Thr Ala Gly Glu Arg Ile Ser Asn Ala Leu
65 70 75 80

Val Ala Met Ala Ile Glu Ser Leu Gly Ala Glu Ala Gln Ser Phe Thr
85 90 95

Gly Ser Gln Ala Gly Val Leu Thr Thr Glu Arg His Gly Asn Ala Arg
100 105 110

Ile Val Asp Val Thr Pro Gly Arg Val Arg Glu Ala Leu Asp Glu Gly
115 120 125

Lys Ile Cys Ile Val Ala Gly Phe Gln Gly Val Asn Lys Glu Thr Arg
130 135 140

Asp Val Thr Thr Leu Gly Arg Gly Gly Ser Asp Thr Thr Ala Val Ala
145 150 155 160

Leu Ala Ala Ala Leu Asn Ala Asp Val Cys Glu Ile Tyr Ser Asp Val
165 170 175

Asp Gly Val Tyr Thr Ala Asp Pro Arg Ile Val Pro Asn Ala Gln Lys
180 185 190

Leu Glu Lys Leu Ser Phe Glu Glu Met Leu Glu Leu Ala Ala Val Gly
195 200 205

Ser Lys Ile Leu Val Leu Arg Ser Val Glu Tyr Ala Arg Ala Phe Asn
210 215 220

Val Pro Leu Arg Val Arg Ser Ser Tyr Ser Asn Asp Pro Gly Thr Leu
225 230 235 240

Ile Ala Gly Ser Met Glu Asp Ile Pro Val Glu Glu Ala Val Leu Thr
245 250 255

Gly Val Ala Thr Asp Lys Ser Glu Ala Lys Val Thr Val Leu Gly Ile
260 265 270

Ser Asp Lys Pro Gly Glu Ala Ala Lys Val Phe Arg Ala Leu Ala Asp
275 280 285

Ala Glu Ile Asn Ile Asp Met Val Leu Gln Asn Val Ser Ser Val Glu
290 295 300

Asp Gly Thr Thr Asp Ile Ile Phe Thr Cys Pro Arg Ser Asp Gly Arg
305 310 315 320

Arg Ala Met Glu Ile Leu Lys Lys Leu Gln Val Gln Gly Asn Trp Thr
325 330 335

Asn Val Leu Tyr Asp Asp Gln Val Gly Lys Val Ser Leu Val Gly Ala
340 345 350

Gly Met Lys Ser His Pro Gly Val Thr Ala Glu Phe Met Glu Ala Leu
355 360 365

Arg Asp Val Asn Val Asn Ile Glu Leu Ile Ser Thr Ser Glu Ile Arg
370 375 380

Ile Ser Val Leu Ile Arg Glu Asp Asp Leu Asp Ala Ala Arg Ala
385 390 395 400

Leu His Glu Gln Phe Gln Leu Gly Gly Glu Asp Glu Ala Val Val Tyr
405 410 415

Ala Gly Thr Gly Arg
420

<210> 66

<211> 5860

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:Plasmid

<400> 66

cccgggtacca cgcgtcccag tggctgagac gcatccgcta aagccccagg aaccctgtgc	60
agaaagaaaa cactcctctg gctaggtaga cacagtttat aaaggtagag ttgagcgggt	120
aactgtcagc acgtagatcg aaaggtgcac aaaggtggcc ctggtcgtac agaaatatgg	180
cggttcctcg cttgagagtg cggaacgcat tagaaacgtc gctgaacgga tcgttgccac	240
caagaaggct ggaaatgatg tcgtggttgt ctgctccgca atgggagaca ccacggatga	300
acttctagaa cttgcagcgg cagtgaatcc cgttccgcca gctcgtgaaa tggatatgct	360
cctgactgct ggtgagcgta tttctaacgc tctcgtcgcc atggctattg agtcccttgg	420

cgcagaagcc caatctttca cgggctctca ggctgggtgtg ctcaccaccg agcgccacgg 480
aaacgcacgc attgttgatg tcactccagg tcgtgtgcgt gaagcactcg atgagggcaa 540
gatctgcatt gttgctgggt tccaggggtg taataaagaa acccgcgatg tcaccacgtt 600
gggtcgtggg ggttctgaca ccaactgcagt tgcgttgga gctgctttga acgctgatgt 660
gtgtgagatt tactcggacg ttgacgggtg gtataccgct gacccgcgca tcgttccctaa 720
tgcacagaag ctggaaaagc tcagcttcga agaaatgctg gaacttgctg ctgttggttc 780
caagattttg gtgctgcgca gtgttgata cgctcgtgca ttcaatgtgc cacttcgcgt 840
acgctcgtct tatagtaatg atccccgac tttgattgcc ggctctatgg aggatattcc 900
tgtggaagaa gcagtcctta ccgggtgtgc aaccgacaag tccgaagcca aagtaaccgt 960
tctgggtatt tccgataagc caggcgaggc tgcgaagggt ttccgtgcgt tggctgatgc 1020
agaaatcaac attgacatgg ttctgcagaa cgtctcttct gtagaagacg gcaccaccga 1080
catcatcttc acctgccctc gttccgacgg ccgcccgcg atggagatct tgaagaagct 1140
tcaggttcag ggcaactgga ccaatgtgct ttacgacgac caggtcggca aagtctccct 1200
cgtgggtgct ggcatgaagt ctcaccacgg tgttaccgca gagttcatgg aagctctgcg 1260
cgatgtcaac gtgaacatcg aattgatttc cacctctgag attcgtattt ccgtgctgat 1320
ccgtgaagat gatctggatg ctgctgcacg tgcattgcat gaggcagttc agctgggcgg 1380
cgaagacgaa gccgtcgttt atgcaggcac cggacgctaa agtttttaaag gagtagtttt 1440
acaatgacca ccatcgcagt tgttggtgca accggccagg tcggccagggt tatgcgccacc 1500
cttttggaag agcgcaattt cccagctgac actgttcgtt tcttttgcttc cccacgttcc 1560
gcaggccgta agattgaatt cgtcgacatc gatgctcttc tgcgttaatt aacaattggg 1620
atcctctaga cccgggattt aaatcgctag cgggctgcta aaggaagcgg aacacgtaga 1680
aagccagtcc gcagaaacgg tgctgacccc ggatgaatgt cagctactgg gctatctgga 1740
caagggaaaa cgcaagcgca aagagaaagc aggtagcttg cagtgggctt acatggcgat 1800
agctagactg ggcggtttta tggacagcaa gcgaaccgga attgccagct ggggcgcctc 1860
ctggtaaggt tgggaagccc tgcaaagtaa actggatggc tttcttgccg ccaaggatct 1920
gatggcgag gggatcaaga tctgatcaag agacaggatg aggatcgttt cgcattgattg 1980
aacaagatgg attgcacgca ggttctccgg ccgcttggtt ggagaggcta ttcggctatg 2040
actgggcaca acagacaatc ggctgctctg atgccgccgt gttccggctg tcagcgcagg 2100
ggcgcccggt tctttttgtc aagaccgacc tgtccgggtg cctgaatgaa ctgcaggacg 2160
aggcagcgcg gctatcgtgg ctggccacga cgggcgttcc ttgcgcagct gtgctcgacg 2220
ttgtcactga agcgggaagg gactggctgc tattgggcga agtgccgggg caggatctcc 2280
tgtcatctca ccttgctcct gccgagaaag tatccatcat ggctgatgca atgcggcggc 2340

tgcatacgct tgatccggct acctgcccac tcgaccacca agcgaaacat cgcacgagc 2400
gagcacgtac tcggatggaa gccgggtcttg tcgatcagga tgatctggac gaagagcatc 2460
aggggctcgc gccagccgaa ctgttcgcca ggctcaaggc gcgcatgccc gacggcgagg 2520
atctctgctg gacccatggc gatgcctgct tgccgaatat catgggtggaa aatggccgct 2580
tttctggatt catcgactgt ggccggctgg gtgtggcgga ccgctatcag gacatagcgt 2640
tggctacccg tgatattgct gaagagcttg gcggcgaaatg ggctgaccgc ttctctgctgc 2700
tttacgggtat cgcgcgtccc gattcgcagc gcacgcctt ctatgcctt cttgacgagt 2760
tcttctgagc gggactctgg gggttcgaaat gaccgaccaa gcgacgcccc acctgccatc 2820
acgagatttc gattccaccg ccgccttcta tgaaagggtg ggcttcggaa tcgttttccg 2880
ggacgccggc tggatgatcc tccagcgcgg ggatctcatg ctggagttct tcgcccacgc 2940
tagcggcgcg ccggccggcc cgggtgtgaa taccgcacag atgcgtaagg agaaaatacc 3000
gcacagggc ctcttccgct tctctgctca ctgactcgt gcgctcggtc gttcggctgc 3060
ggcgagcggc atcagctcac tcaaaggcgg taatacgggt atccacagaa tcaggggata 3120
acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg 3180
cgttgctggc gtttttccat aggtcgcgc cccctgacga gcacacaaa aatcgacgt 3240
caagtcagag gtggcgaaac ccgacaggac tataaagata ccaggcgttt cccctggaa 3300
gtccctcgt gcgctctct gttccgacce tgccgcttac cggatacctg tccgcctttc 3360
tcccttcggg aagcgtggcg ctttctcata gctcacgctg taggtatctc agttcgggtg 3420
aggctcgttcg ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg 3480
ccttatccgg taactatcgt cttgagtcca acccggtgta acacgactta tcgccactgg 3540
cagcagccac tggtaacagg attagcagag cgaggatatg aggcgggtgct acagagttct 3600
tgaagtggcg gcctaactac ggctacacta gaaggacagt atttggtatc tgcgctctgc 3660
tgaagccagt taccttcgga aaaagagttg gtagctcttg atccggcaaa caaaccaccg 3720
ctggtagcgg tgggtttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc 3780
aagaagatcc tttgatcttt tctacggggt ctgacgctca gtggaacgaa aactcacgtt 3840
aagggtttt ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt ttaaaggccg 3900
gccgcggccg ccacggcat tttcttttgc gtttttattt gttaactgtt aattgtcctt 3960
gttcaaggat gctgtctttg acaacagatg ttttcttgcc tttgatgttc agcaggaagc 4020
tcggcgcaaa cgttgattgt ttgtctcgt agaatcctct gtttgtcata tagcttgtaa 4080
tcacgacatt gtttcctttc gcttgaggta cagcgaagtg tgagtaagta aaggttacat 4140
cgttaggatc aagatccatt tttaacacaa ggccagttt gttcagcggc ttgtatgggc 4200

cagttaaaga attagaaaca taaccaagca tgtaaataatc gtttagacgta atgccgtcaa 4260
tcgtcatttt tgatccgcgg gagtcagtga acaggtacca ttgccgttc attttaaaga 4320
cgttcgcgcg ttcaatttca tctgttactg tgtttagatgc aatcagcggg ttcatcactt 4380
ttttcagtgt gtaatcatcg ttttagctcaa tcataccgag agcgcggttt gctaactcag 4440
ccgtgcggtt ttatcgctt tgcagaagtt tttgactttc ttgacggaag aatgatgtgc 4500
ttttgccata gtatgctttg ttaaataaag attcttcgcc ttggtagcca tcttcagttc 4560
cagtgtttgc ttcaaatact aagtatttgt ggccctttatc ttctacgtag tgaggatctc 4620
tcagcgtatg gttgtcgctt gagctgtagt tgccttcacg gatgaactgc tgtacatttt 4680
gatacgtttt tccgtcaccg tcaaagattg atttataatc ctctacaccg ttgatgttca 4740
aagagctgtc tgatgctgat acgttaactt gtgcagttgt cagtgtttgt ttgccgtaat 4800
gtttaccgga gaaatcagtg tagaataaac ggatttttcc gtcagatgta aatgtggctg 4860
aacctgacca ttcttgtgtt tggcttttta ggatagaatc atttgcacg aatttgctgc 4920
tgtctttaaa gacgcggcca gcgtttttcc agctgtcaat agaagtttcg ccgacttttt 4980
gatagaacat gtaaatacgat gtgtcatccg catttttagg atctccggct aatgcaaaga 5040
cgatgtggta gccgtgatag tttgcgacag tgcgcgcagc gttttgtaat ggccagctgt 5100
cccaaacgtc caggcctttt gcagaagaga tatttttaat tgtggacgaa tcaaattcag 5160
aaacttgata tttttcattt ttttgctgtt cagggatttg cagcatatca tggcgtgtaa 5220
tatgggaaat gccgtatgtt tccttatatg gcttttggtt cgtttctttc gcaaacgctt 5280
gagttgcgcc tcctgccagc agtgcggtag taaagggtta tactgttgct tgttttgcaa 5340
actttttgat gttcatcggt catgtctcct tttttatgta ctgtgttagc ggtctgcttc 5400
ttccagccct cctgtttgaa gatggcaagt tagttacgca caataaaaaa agacctaaaa 5460
tatgtaaggg gtgacgcaa agtatacact ttgcccttta cacatttttag gtcttgccgtg 5520
ctttatcagt aacaaacccg cgcgatttac ttttcgacct cattctatta gactctcggtt 5580
tggattgcaa ctggtctatt ttctctttt gtttgataga aaatcataaa aggatttgca 5640
gactacgggc ctaaagaact aaaaaatcta tctgtttctt ttcattctct gtatttttta 5700
tagtttctgt tgcattgggca taaagttgcc tttttaatca caattcagaa aatatcataa 5760
tatctcattt cactaaataa tagtgaacgg caggtatatg tgatgggtta aaaaggatcg 5820
gcgccgcgtc gatttaaatac tcgagaggcc tgacgtcggg 5860

<210> 67

<211> 29

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 67

gagactcgag gttggctggt catcatagg

29

<210> 68

<211> 34

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:PCR primer

<400> 68

gaagagagca tatgtcagcg ctctagtttg gttc

34

<210> 69

<211> 6472

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:Plasmid

<400> 69

tcgagggttg ctggtcatca taggaatcaa cctggccact ttatggtggg caccaccgtc	60
gcaaacaaca tatcttgag caggcgtgtc gattctttcc gccatcattg tttggtttct	120
tcccggcgca cccccgtat ggaatcgccg tcgcattgct tcacgcaaac aacagtccac	180
cggtagacgt cgacaagccc ccaaacgac aagccaccct caaacggcgg aatttagcca	240
acaacaatag actagacaga gctgtccatg tagcatgaac tcgattatca actgccacga	300
gaggtcgggg tcatgtcac caccacaggg acgtcacgc accaaaaaat cggagacttt	360
tacaccgaag ccggagcgac gttcacgac gtaaccatcg cctaccaagc atggggccac	420
tacaccggca ccaatctcat cgttctcgaa catgccctga ccggcgactc taacgctatt	480
tcatggtggg acggactgat tggccctggc aaagcactcg acaccaaccg ctactgcac	540
ctatgcacca acgtgctcgg aggatgcaaa ggatccaccg gaccgagcag tccacacca	600
gacggaaaac catggggatc cagatttcca gccctttcaa tccgtgacct tgtcaatgcc	660
gaaaaacaac ttttcgacca cctcggcatc aataaaattc acgcaatcat cggcggatcc	720
atggggaggcg cagcaccct cgaatgggct gcactccacc cacacatgat gacgactgga	780
ttcgtcatag cagtctcagc acgcgcaagc gcttggcaaa tcggtattca aactgcacaa	840
atcagcgcca tagaactcga cccccactgg aacggcggcg attactacag cggtcacgca	900
ccatgggaag gaatcgccgc cgctcgccgg atcgcccacc tcacctatcg cggcgaacta	960
gaaatagacg aacgattcgg cacttcgcga caacacggtg aaaaccact cggccccctt	1020
cgagatccac atcaacgttt tgcggtcacg agctacctcc aacaccaagg catcaaactc	1080

gctcaacgat tcgatgcagg tagttacgtc gtgcttaccg aagccctcaa tcgtcatgac 1140
atcggacgcg gccgaggcgg actcaacaaa gccctcagcg caatcacagt ccccatcatg 1200
attgctggcg ttgataccga tattctctac ccctatcacc agcaagaaca cctatcacga 1260
aatctaggca acctactcgc tatggcaaaa atcagctcac cagtaggcca cgacgctttc 1320
ctcacagaat tccgacaaat ggagcgaatc ctaagacatt tcatggagct ttcggaagga 1380
atcgacgatt ccttccgaac caaactagag cgctgacata tgactagttc ggacctaggg 1440
atatcgtcga catcgatgct cttctgcgtt aattaacaat tgggatcctc tagaccggg 1500
atttaaactc ctagcgggct gctaaaggaa gcggaacacg tagaaagcca gtccgcagaa 1560
acgggtgctga ccccgatga atgtcagcta ctgggctatc tggacaaggg aaaacgcaag 1620
cgcaaagaga aagcaggtag cttgcagtg gcttacatgg cgatagctag actgggcggt 1680
tttatggaca gcaagcgaac cggaattgcc agctggggcg ccctctggta aggttgggaa 1740
gccctgcaa gtaaactgga tggctttctt gccgccagg atctgatggc gcaggggatc 1800
aagatctgat caagagacag gatgaggatc gtttcgcatg attgaacaag atggattgca 1860
cgcaggttct ccggccgctt gggtaggag gctattcggc tatgactggg cacaacagac 1920
aatcggctgc tctgatgccg ccgtgttccg gctgtcagcg caggggcgcc cggttctttt 1980
tgtcaagacc gacctgtccg gtgccctgaa tgaactgcag gacgaggcag cgcggctatc 2040
gtggctggcc acgacgggcg ttccttgccg agctgtgctc gacgttgtca ctgaagcggg 2100
aagggactgg ctgctattgg gcgaagtgcc ggggcaggat ctctgtcat ctcaccttgc 2160
tcctgccgag aaagtatcca tcatggctga tgcaatgcgg cggctgcata cgcttgatcc 2220
ggctacctgc ccattcgacc accaagcgaa acatcgcatc gagcgagcac gtactcggat 2280
ggaagccggt cttgtcgatc aggatgatct ggacgaagag catcaggggc tcgcgccagc 2340
cgaactgttc gccaggctca aggcgcgcac gcccgacggc gaggatctcg tcgtgaccca 2400
tggcgatgcc tgcttgccga atatcatggg ggaaaatggc cgcttttctg gattcatcga 2460
ctgtggccgg ctgggtgtgg cggaccgcta tcaggacata gcgttggtta cccgtgatat 2520
tgctgaagag cttggcggcg aatgggctga ccgcttctc gtgctttacg gtatcgccgc 2580
tcccgatctg cagcgcacatc cttctatctg cttcttgac gagttcttct gagcgggact 2640
ctggggttcg aaatgaccga ccaagcgacg cccaacctgc catcacgaga ttctgatcc 2700
accgccgcct tctatgaaag gttgggcttc ggaatcgttt tccgggacgc cggctggatg 2760
atcctccagc gcggggatct catgctggag ttcttcgccc acgctagcgg cgcgccggcc 2820
ggcccggtgt gaaataccgc acagatgcgt aaggagaaaa taccgcatca ggcgctcttc 2880
cgcttctcgc ctcactgact cgctgcgctc ggtcgttcgg ctgcggcgag cggtatcagc 2940
tcactcaaag gcggtaatac ggttatccac agaatacagg gataacgcag gaaagaacat 3000

gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt 3060
ccataggctc cgccccctg acgagcatca caaaaatcga cgctcaagtc agaggtggcg 3120
aaacccgaca ggactataaa gataccaggc gtttccccct ggaagctccc tcgtgcgctc 3180
tcctgttccg accctgccgc ttaccggata cctgtccgcc tttctccctt cgggaagcgt 3240
ggcgctttct catagctcac gctgtaggta tctcagttcg gtgtaggtcg ttcgctccaa 3300
gctgggctgt gtgcacgaac cccccgttca gcccgaccgc tgcgccttat ccggtaaacta 3360
tcgtcttgag tccaacccgg taagacacga cttatcgcca ctggcagcag ccactggtaa 3420
caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa 3480
ctacggctac actagaagga cagtatttgg tatctgcgct ctgctgaagc cagttacctt 3540
cggaaaaaga gttggtagct cttgatccgg caaacaacc accgctggta gcggtggttt 3600
ttttgtttgc aagcagcaga ttacgcgcag aaaaaaagga tctcaagaag atcctttgat 3660
cttttctacg ggggtctgacg ctcagtggaa cgaaaactca cgttaaggga ttttggctat 3720
gagattatca aaaaggatct tcacctagat ctttttaaag gccggccgcg gccgcgcaaa 3780
gtcccgttcc gtgaaaattt tcgtgccgcg tgattttccg ccaaaaactt taacgaacgt 3840
tcgttataat ggtgtcatga ctttcacgac gaagtactaa aattggcccc aatcatcagc 3900
tatggatctc tctgatgtcg cgctggagtc cgacgcgctc gatgctgccg tcgatttaaa 3960
aacggtgatc ggatttttcc gagctctcga tacgacggac gcgccagcat cacgagactg 4020
ggccagtgcc gcgagcgacc tagaaactct cgtggcggat cttgaggagc tggctgacga 4080
gctgcgtgct cggccagcgc caggaggacg cacagtagtg gaggatgcaa tcagttgcgc 4140
ctactgcggt ggctgatctc ctccccggcc tgaccgcgca ggacggcgcg caaaatattg 4200
ctcagatgcg tgtcgtgccg cagccagccg cgagcgcgcc aacaaacgcc acgccgagga 4260
gctggaggcg gctaggtcgc aaatggcgct ggaagtgcgt cccccgagcg aaattttggc 4320
catggtcgtc acagagctgg aagcggcagc gagaattatc gcgatcgtgg cggtgcccg 4380
aggcatgaca aacatcgtaa atgccgcgtt tcgtgtgccg tggccgcccc ggacgtgtca 4440
gcgccgccac cacctgcacc gaatcggcag cagcgtcgcg cgtcgaaaaa gcgcacaggc 4500
ggcaagaagc gataagctgc acgaatacct gaaaaatgtt gaacgccccg tgagcggtaa 4560
ctcacagggc gtcggctaac ccccagtcca aacctgggag aaagcgctca aaaatgactc 4620
tagcggattc acgagacatt gacacaccgg cctggaaatt ttccgctgat ctgttcgaca 4680
cccatcccga gctcgcgctg cgatcacgtg gctggacgag cgaagaccgc cgcgaattcc 4740
tcgctcacct gggcagagaa aatttccagg gcagcaagac ccgcgacttc gccagcgctt 4800
ggatcaaaga cccggacacg gagaaacaca gccgaagtta taccgagttg gttcaaaatc 4860

gcttgcccgg tgccagtatg ttgctctgac gcacgcgcag cacgcagccg tgcttgtcct 4920
ggacattgat gtgccgagcc accaggcccg cgggaaaatc gagcacgtaa accccgaggt 4980
ctacgcgatt ttggagcgct gggcacgcct ggaaaaagcg ccagcttgga tcggcgtgaa 5040
tccactgagc gggaaatgcc agctcatctg gctcattgat ccggtgtatg ccgcagcagg 5100
catgagcagc ccgaatatgc gcctgctggc tgcaacgacc gaggaaatga cccgcgtttt 5160
cggcgctgac caggcttttt cacataggct gagccgtggc cactgcactc tccgacgac 5220
ccagccgtac cgctggcatg ccagcacaa tcgctgggat cgcttagctg atcttatgga 5280
ggttgctcgc atgatctcag gcacagaaaa acctaaaaaa cgctatgagc aggagttttc 5340
tagcggacgg gcacgtatcg aagcggcaag aaaagccact gcggaagcaa aagcacttgc 5400
cacgcttgaa gcaagcctgc cgagcgccgc tgaagcgtct ggagagctga tcgacggcgt 5460
ccgtgtcctc tggactgctc cagggcgtgc cgcccgatg gagacggctt ttcgccacgc 5520
tttgactgtg ggataccagt taaaagcggc tggtagcgc ctaaaagaca ccaagggta 5580
tcgagcctac gagcgtgcct acaccgtgc tcaggcggtc ggaggaggcc gtgagcctga 5640
tctgccgccg gactgtgacc gccagacgga ttggccgcga cgtgtgcgcg gctacgtcgc 5700
taaaggccag ccagtcgtcc ctgctcgtca gacagagacg cagagccagc cgaggcgaaa 5760
agctctggcc actatgggaa gacgtggcgg taaaaggcc gcagaacgct ggaaagaccc 5820
aaacagtgag tacgcccag cacagcgaga aaaactagct aagtcagtc aacgacaagc 5880
taggaaagct aaaggaaatc gcttgaccat tgcaggttgg ttatgactg ttgagggaga 5940
gactggctcg tggccgacaa tcaatgaagc tatgtctgaa tttagcgtgt cacgtcagac 6000
cgtgaataga gcacttaagg tctgcgggca ttgaacttcc acgaggacgc cgaaagcttc 6060
ccagtaaatg tgccatctcg taggcagaaa acggttcccc cgtagggctc ctctcttggc 6120
ctcctttcta ggtcgggctg attgctcttg aagctctcta ggggggctca caccataggc 6180
agataacgtt cccaccggc tcgcctcgta agcgacaaag gactgctccc aaagatcttc 6240
aaagccactg ccgcgactgc cttcgcgaag ccttgccccg cggaatttc ctccaccgag 6300
ttcgtgcaca cccctatgcc aagcttcttt caccctaaat tcgagagatt ggattcttac 6360
cgtggaaatt cttcgcaaaa atcgtccctt gatcgccctt gcgacgttgg cgtcggtgcc 6420
gctggttgcg cttggcttga ccgacttgat cagcggccgc tcgatttaaa tc 6472